DOCUMENT RESUME

ED 052 166 24 SP 005 102

TITLE	A Proposal Design for Training Educational Evaluators and Developers. Technical Paper No. 3. Final Report.
INSTITUTION	Colorado Univ., Boulder. Colorado Center for Training in Educational Evaluation and Development.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
BUREAU NO	BR-0-9034
CONTRACT	OEC-0-70-4770(520)
NOTE	105p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	*Consortia, *Educational Development, *Educational Programs, *Evaluation Methods, *Teacher Education

ABSTRACT

In this third technical paper, the design for a training program for educational developers and evaluators is explicated. Mentioned are 1) the procedures by which the design was generated and 2) the salient characteristics of the several agencies which make up the consortium. The topics considered in this paper include 1) a general description of the highlights of the proposed training; 2) the basic assumptions underlying the training program; 3) the objectives of the training program; 4) the elements of the three-level training program (Ph.D. program, MA program, and intensive training institutes); 5) the staff who will be associated with the program; 6) the recruitment and selection of trainees; 7) the materials development proposed; 8) the evaluation plan for the training program; and 9) the administrative, monitoring, and cost accounting elements of the training system. [Related documents are SP 005 100-101, and 005 103.] (MBM)

ED052166

BR 0-9034
PA 24
SP

Final Report

Technical Paper Number 3

A PROPOSED DESIGN FOR TRAINING EDUCATIONAL EVALUATORS AND DEVELOPERS

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Colorado Center for Training in
Educational Evaluation and Development
University of Colorado
Boulder, Colorado
80302

SP005102

The work reported herein was produced under contract with the U.S. Office of Education, Contract No. OEC-0-70-4770(520).

TECHNICAL PAPER NUMBER 3

A PROPOSED DESIGN FOR TRAINING EDUCATIONAL EVALUATORS AND DEVELOPERS

TABLE OF CONTENTS

		Page
I.	General Description of the Training Program	2
II.	General Assumptions Underlying the Training Program	4
III.	Objectives of the Training Program	17
IV.	Elements of the Training Program	33
	A. Graduate Level Training in Evaluation and Development . . .	33
	B. Intensive Training Institutes	55
V.	Staff of the Training Program	62
VI.	Recruitment and Selection of Trainees	69
VII.	Proposed Materials Development	75
VIII.	Evaluation of the Training Programs	83
IX.	Administering, Monitoring, and Determining the Cost Effectiveness of the Training Program	96

Appendix A. Vitae of Staff.

Appendix B. Progress Report and Evaluation of the Graduate Research Training Program and the Laboratory of Educational Research.

Appendix C. Sample Measurement Instruments.

Appendix D. Evaluation of Long-Range Effects of a 1967 Research Training Presession.

A PROPOSED DESIGN FOR TRAINING EDUCATIONAL EVALUATORS AND DEVELOPERS

In this third technical paper, the design for a training program for educational developers and evaluators is explicated. Mentioned only in passing are (1) the procedures by which the design was generated and (2) the salient characteristics of the several agencies that make up the consortium; the reader interested in these elements is referred to Technical Papers Numbers 1 and 2, respectively, in this series.

The topics considered in this section include the following:

1) A general description of the highlights of the proposed training.
2) The basic assumptions underlying the training program.
3) The objectives of the training program.
4) The elements of the three-level training program (Ph.D program, MA program, and intensive training institutes).
5) The staff who will be associated with the program.
6) The recruitment and selection of trainees.
7) The materials development proposed.
8) The evaluation plan for the training program.
9) The administrative, monitoring, and cost accounting elements of the training system.

I. General Description of the Training Program.

Generally, a multi-faceted training program is being proposed. Two elements concern training leading to graduate level degrees at both the masters and doctorate level. A third element involves intensive training institutes, primarily for personnel already on-the-job. The training program involves over a dozen consortium agencies, as was detailed in Technical Paper Number 2.

Training leading to graduate degrees is proposed in four areas: the masters degree in evaluation; the masters degree in development; the doctorate in development; and the doctorate in evaluation. Specific content and planned experiences for each of these degrees is explicated in Section IV below. These activities include a wide variety of exceptional internship opportunities. Recruitment will be national in scope, and screening will be intensive (see Section VI below).

The intensive training institutes in evaluation and development will be held in increments of short time duration (usually varying in length between one and four weeks). They will be individually designed and tailored to meet the needs of the particular group of trainees that is coming for the training (see Section IV). An experiment to determine the effectiveness of various follow-up models is being proposed (Section VIII). In regard to the training program, especially the short-term institutes, considerable materials development is contemplated (Section VII).

It can be noted in the written documents associated with designing the project that there has been an extensive shift in orientation over the six-month design period. This shift might best be described as away

from the traditional university-oriented model to a user-oriented model, with elaborate provisions for meaningful internship experiences for pre-service trainees and high-powered short-term institutes for two types of personnel: those already holding positions with evaluation and development responsibilities and those whose responsibilities are in the process of being redefined and will include development and evaluation. Additonally, many meaningful and communicative bridges are contemplated linking the various training programs proposed under this project, as well as linking them to existing training programs at Colorado and at many of the consortium units.

All elements of the training program designed are considered in detail in the sections that follow in this technical paper.

II. General Assumptions Underlying the Training Program

Below are discussed several assumptions, each of which has had a direct influence on the design of the proposed training programs. Supporting evidence for the assumptions ranges from hunches to relatively hard data and is indicated as each assumption is presented. It is encouraging to the developers of the design that each assumption does appear to be supportable, at least in large part.

1) There is a need for pre-service, graduate level training in evaluation and development.

Supporting this assumption are a number of extant manpower studies. Most recent and relevant for the design project have been two studies completed by the AERA Task Force on Research Training. The first concerned analyses of AERA placement data for the years 1968, 1969, and 1970 (referred to initially in Technical Paper Number 1 in this series). A definite need for new entrants into the fields of evaluation and development was apparent over that time period, although it was somewhat less pronounced recently with the advent of constricted federal funding. The second AERA task force study of fifty-eight employers of R, D, D, and E personnel resulted in similar findings.

Information collected during the course of the design grant confirms this need. Most institutions in this consortium, probably representing less than one percent of the potential employment spectrum in R, D, D, and E, expressed needs for additional personnel in evaluation and development, especially at the master's level. Possibly even a greater need identified was the need for additional training for present employees (see the working papers of the consortium units, Appendices A - G, in Technical Paper Number 2).

Data collected via questionnaire (slightly over fifty percent returns) resulted in compatible indications (the actual questionnaire was presented in Appendix B in Technical Paper Number 1). Results for evaluation are presented in Table 1; in Table 2, development data are presented.* The tables are self-explanatory; and trends are apparent. Even with no funding increase, about forty percent of the agencies intend to hire staff in both evaluation and development. With a twenty percent funding increase, the intention-to-hire figure jumps to over eighty percent of the organizations. Note, also, that several types of agencies plan to hire more than one person at each level, and that the greatest demand seems to be at the master's level (for both evaluators and developers).

2) There does not exist, at this time, even a moderate number of institutions that provide training in development.

Project staff had inclinations to believe this assumption even before the design project commenced. As part of the project, it was possible to talk to many persons from centers and regional laboratories. They reinforced this opinion. The national curriculum projects also made it clear in their working papers (see appendices in Technical Paper Number 2) that there is no single training institution to which they go when hiring new staff. Rather, they determine the skills of the particular role assignment that they hope to fill and then conduct an elaborate search, often nationwide, to find such a person.

---

* As over forty percent of the sampled agencies did not respond, the percentages in Tables 1, 2, and 3 may be inflated and should be interpreted with caution. That is, many organizations electing not to return the questionnaire may have considered the training irrelevant and/or not feasible.

Table 1. Percent of organizations intending to hire evaluators at the BS, MS, and PhD levels with and without funding increase.

Type of Organization	With 20% budget increase, percent of organizations that would hire evaluators at _____ level.				With no budget increase, percent of organizations that would hire evaluators at _____ level.			
	Some	BS	MS	PhD	Some	BS	MS	PhD
Labs, Centers, Nat'l Curr. Projects	91%	18% (3.5)	45% (1.3)	73% (1.3)	56%	18% (2.0)	27% (1.2)	45% (0.9)
Major School Districts [a]	88%	38% (2.7)	88% (2.7)	75% (1.7)	88%	25% (1.5)	75% (2.0)	38% (1.0)
Intermediate School Districts [b]	83%	6% (4.0)	67% (2.0)	22% (1.0)	6%	0 (0)	0 (0)	6% (1.0)
State Education Departments	91%	26% (1.4)	82% (2.0)	50% (1.6)	41%	5% (1.0)	36% (1.1)	19% (1.3)
Miscellaneous Agencies [c]	79%	14% (1.3)	64% (1.3)	43% (1.0)	29%	0 (0)	14% (1.8)	14% (1.0)
Average	86%	16% (2.4) [d]	70% (1.9)	48% (1.4)	37%	7% (1.6)	26% (1.5)	21% (1.0)

[a] Major school districts are the districts in the 30 largest cities in the U. S. A.

[b] Intermediate school districts are defined as those having 2,500 to 10,000 students.

[c] Miscellaneous agencies include foundations, private research organizations, test publishers, etc.

[d] Numbers in parentheses indicate the average number who would be hired at that level. In this case, of the agencies intending to hire evaluators at the BS level, each plans to hire 2.4, on the average.

Table 2. Percent of organizations intending to hire developers at the BS, MS, and PhD levels with and without funding increase.

Type of Organization	With 20% budget increase, percent of organizations that would hire developers at _____ level.				With no budget increase, percent of organizations that would hire developers at _____ level.			
	Some	BS	MS	PhD	Some	BS	MS	PhD
Labs, Centers, Nat'l Curr. Projects	91%	45% (5.8)	82% (3.1)	82% (2.2)	55%	36% (3.8)	36% (1.8)	45% (1.6)
Major School Districts [a]	75%	38% (2.3)	75% (3.2)	38% (1.3)	75%	38% (1.3)	63% (2,2)	25% (1.5)
Intermediate School Districts [b]	83%	17% (1.0)	72% (2.7)	22% (1.0)	17%	0 (0)	17% (3.2)	0 (0)
State Education Departments	82%	18% (1.0)	65% (1.9)	41% (1.3)	33%	6% (1.0)	29% (1.4)	12% (1.0)
Miscellaneous Agencies [c]	79%	14% (1.5)	57% (2.1)	43% (2.0)	43%	7% (1.0)	21% (1.3)	36% (1.0)
Average	82%	24% (2.8) [d]	69% (2.5)	43% (1.7)	40%	13% (2.3)	29% (1.9)	21% (1.3)

[a] Major school districts are the districts in the 30 largest cities in the U. S. A.

[b] Intermediate school districts are defined as those having 2,500 to 10,000 students.

[c] Miscellaneous agencies include foundations, private research organizations, test publishers, etc.

[d] Numbers in parentheses indicate the average number who would be hired at that level. In this case, of the agencies intending to hire developers at the BS level, each plans to hire 2.8, on the average.

It also can be noted in passing that there is only a limited number of organizations that claim to have formal training programs in evaluation. In the past, many persons graduating from research training programs have picked up assorted evaluation skills in the process. The various conceptual papers recently written should facilitate the process of detailing the specifics of the research, development, diffusion, and evaluation domains; it is possible that much done previously in the name of "research" actually pertained more to evaluation than to research.

3) There is a jumble of literature dealing with diffusion and change agents, but the existence of hard data in this area is rare.

Like many of the other design projects, personnel connected with this consortium made a valiant effort to conceptualize what training in diffusion would be like. There is no question but that certain of the skills considered important, particularly by personnel in the national curriculum projects, clearly are critical for diffusion. Indeed, the relationships between development, evaluation, and diffusion appear to be extensive.

Be that as it may, it is extremely difficult to locate and study data dealing with diffusion, let alone extracting from the data the principles upon which a training program for diffusers might be based. Although the literature on such roles as change agent, diffuser, and the like is abundant, evidence and specifics about effective methodologies for them to use and critical ingredients of their training are essentially unknown. Being blessed with normal vision, the

preparers of this proposal were aware of the vacuum existing in the diffusion area and recognized the natural inclination to rush into that uninhabited space in order to increase the likelihood of funding. The project staff and the consortium agencies, in this case, have successfully defended against that inclination; there are no plans that this program train diffusers. Naturally, there are diffusion-related skills that are a part of evaluation and especially, development. Examples of such skills are the use of media, the use of ERIC, knowledge of professional organizations and modes of communication they use (e.g., annual conventions), and the ability to communicate in writing. In the programs proposed herein, these activities and skills are viewed as supporting, rather than central, in nature.

Therefore, in this program, training in diffusion skills will be incidental rather than planned. It should be noted, however, that our inability to make advancements in the diffusion area will make us doubly alert and prone to examine closely those designs that propose to train diffusers. Such efforts should be scrutinized to ascertain either the brilliance of the designers or their effrontery.

4) There is a need for intensive training institutes in evaluation and development for personnel whose roles call for skills in these areas, or for those whose roles likely will be redefined as a result of such training.

Extensive input suggesting that this was a valid assumption first occurred in the working papers submitted by the consortium units (see appendices to Technical Paper Number 2) and also from our discussions with consortium agencies that did not submit working papers. In part probably explained by the current financial squeeze, many agencies were looking "to make do" with their present personnel.

As a result, they have pronounced needs for various types of on-the-job or off-the-job training to either develop new skills in employees or refurbish existing skills in certain evaluation and development areas. It also can be noted that job redefinition is needed in many cases. Institutes such as those planned should add to the pool of trained manpower in evaluation and development. The widespread consensus on the need for training institutes has been relected in the substantial level of resources proposed for that purpose in this design; institute training clearly represents a substantial departure from what was inferred in our original response to RFP 70-12.

Responses on the questionnaires that were sent out reinforced our feelings in this regard; data are presented in Table 3. It can be noted that approximately 95 percent of the agencies returning questionnaires indicated interest in sending present employees to evaluation institutes; the corresponding figure for development institutes was 96 percent. The averages presented in the table highlight the fact that most agencies were planning to send more than a single staff member. Recall that organizations would continue to pay the individual his regular salary while he attended the institute. This pronounced interest on the part of user organizations is responded to, and reflected in, the proposed training program. Particularly encouraging is that fact that the institutes, if conducted at an intense enough level, can serve a function similar to bringing new entrants into the fields of evaluation and development. That is, the skills imparted to trainees will allow them to undertake, quite possibly for the first time, responsibilities that are heavily evaluative

Table 3. Percent of organizations indicating willingness to send employees to evaluation and development institutes of varying duration.

Type of organization	Percent of organizations that would send employees to ____ weeks of evaluation institutes per year.				Percent of organizations that would send employees to ____ weeks of development institutes per year.			
	Some	Two	Four	Six	Some	Two	Four	Six
Labs, Centers, Nat'l Curr. Projects	82%	82% (3.4)	36% (2.5)	18% (2.0)	91%	91% (6.0)	45% (2.8)	18% (2.0)
Major School Districts[a]	100%	100% (6.3)	50% (4.3)	50% (3.3)	100%	100% (6.9)	63% (5.0)	50% (4.3)
Intermediate School Districts[b]	94%	94% (4.6)	39% (1.9)	33% (1.5)	100%	100% (4.4)	39% (4.0)	39% (3.1)
State Education Departments	93%	93% (3.6)	67% (1.7)	60% (1.5)	94%	94% (3.4)	71% (1.7)	53% (1.2)
Miscellaneous Agencies[c]	100%	100% (3.0)	57% (1.9)	21% (1.3)	93%	93% (1.7)	57% (1.3)	7% (1.0)
Average	95%	95% (4.0)[d]	53% (2.1)	36% (1.8)	96%	96% (4.2)	54% (2.6)	34% (2.4)

[a] Major school districts are the districts in the 30 largest cities in the U. S. A.

[b] Intermediate school districts are defined as those having 2,500 to 10,000 students.

[c] Miscellaneous agencies include foundations, private research organizations, test publishers, etc.

[d] Numbers in parentheses indicate the average number who would be permitted to attend the institute. In this case, of the organizations allowing employees to attend, each plans to let 4.0 take part, on the average.

and developmental in nature; their job descriptions can be altered to reflect their new capabilities and, in effect, they become part of the E and D manpower pool.

5) The length of planned institutes will, to a great degree, serve as a powerful determinant of the number of persons able to take part in the training.

This assumption was formulated primarily on the basis of questionnaire returns. Many organizations were reluctant to release their employees for longer than two weeks. It can be noted in Table 3 that the figures of 95 percent and 96 percent, for evaluation and development institutes, respectively, are for an institute two weeks in length. The percentages fall off from that point; the figure becomes about fifty percent for two two-week institutes (a commitment of four weeks during the year) and drops to approximately thirty-five percent when the commitment becomes six weeks during the year. It would appear that most groups could not attend institutes in excess of four weeks during a single year, and that employees in certain categories might be able to attend only two weeks per year. From comments written in on the questionnaire, it is also apparent that timing of the institutes is critical; there are certain peak work periods during the year when many potential trainees would not be able to be released from their present employment.

6) It is assumed that pronounced changes must take place in the context in which, and the contingencies under which, persons trained in the institutes are employed in order that their new skills can be effectively utilized.

The assumption being made here, and one for which we have only the most subjective type of data (primarily informal observations),

is that it is not simply the lack of skills in development and evaluation that is apparent in most of the educational organizations in this country. Rather the assumption is forwarded that there are well-defined and powerful inhibitory forces that keep even well-learned skills from reaching fruitful expression in educational settings. As an example, consider the frustrated director of research or the recently-trained Title I evaluator who finds that his job description and responsibilities in no way reflect the training he has received; there is likely a gulf between the perceptions of a superintendent and an evaluator as to what the evaluator's role and responsibilities are.

In other words, the assumption is that intensive institutes that further develop and perfect evaluation and development skills still will not be sufficient to insure that persons thus trained will function effectively in their home environment. Thus, the design staff believes that it is necessary and important to instruct other personnel in school and educational administrative hierarchies (1) on the proper role and effective use of decision-oriented evaluators and (2) on the proper and effective use of developers. It is proposed that such supplemental instruction be undertaken as one phase of the operation of the training program; it would appear to be a vital service function.

7) The assumption is made that research is needed to determine the effectiveness of various possible follow-up models at the con-conclusion of institute training.

It is normally assumed, and probably rightly so, that even intensive institute training does not have long-range substantial effects

unless it is accompanied by follow-up experiences. Obviously, varying degrees of follow-up could be undertaken; variations in intensity are normally matched by concomminant variations in the amount of resources required for the follow-up. It seems reasonable that there exists, somewhere among the infinite variations of possible follow-up experiences, an optimal procedure, or at least procedures that are most cost effective.

This being the case, it is proposed in the evaluation section of this proposal (Section VIII) that questions of this nature be answered experimentally. A random group of trainees would be given extensive follow-up on the job, a second group would be called back to the training site for additional sessions, a third group would be worked with only by phone and mail, and a final group would receive the traditional treatment (that is, essentially no follow-up at all). It is felt that such a procedure, if used in the first eighteen months or so of the training, would permit the answering of questions that would then allow structuring around the most effective follow-up procedure during the remainder of the operational period of the training program.

8) It is assumed that there exists extensive needs for the development of training materials in the areas of evaluation and development.

A recent RFP from the U.S. Office of Education requested new ideas for training material packages. The AERA task force has likewise conducted a search for promising training practices and materials. All indications point to a dearth of effective materials; possibly more important, there seems to be a pronounced shortage of viable training alternatives or novel ideas in the area. In this particular

program, needs are felt primarily in two areas. First, a need exists for simulation materials that might be used effectively with training in evaluation and development. A second apparent need is for the development of training package modules for the model of intensive training institutes proposed. The need for simulation materials and materials of the latter type, the training package modules, is elaborated upon in Section VII. of this technical paper. Also, by referring to Section IV.B. of this paper, the reader can get a better idea of how the training modules will permit, to a large extent, individualizing and tailoring the training institutes to the specific needs of particular groups.

9) It is assumed that the training program design, though innovative, must also be feasible.

A very real trade off exists when one tries to design a training program that abounds with originality and novelty, and yet at the same time must be feasible enough that it has substantial operational validity. The Colorado consortium has attempted to proceed at the fulcrum between these two extremes. It clearly is difficult to break out of existing training modes that have existed for several decades in the schools and universities. It is also difficult to propose changes so sweeping that their feasibility and practicality are questioned by the essentially conservative educational establishment. Nevertheless, it is felt that a reasonable balance has been reached between the need for innovativeness and the need for feasibility. Further, it is recognized that these two qualities are not incompatible within the same training program.

The assumptions outlined in this section were influential in effecting the training program that was designed. The assumptions plus multiple sources of data input led to the establishment of numerous objectives for the training program, and these objectives are explicated in the following section of this paper.

III. Objectives of the Training Program.

Goals for the training programs can be stated at at least three levels of generality. At the most general level are broad goals of training which relate to the types of person to be recruited, a general description of the activities in which they will participate, and job descriptions and descriptions of work settings for trainees. At a greater level of specificity are complexes of skills which are to be imparted through the training program. At the most specific level are behavioral statements of instructional objectives which guide the teaching-learning process on a day-to-day basis.

Broad Goals

At this level of generality we do not distinguish between goals for training evaluators and goals for training development specialists. The broad goals of the program are to attract competent individuals into the areas of educational evaluation and development and train them to assume responsible roles as evaluators and developers in research and development centers, regional laboratories, school districts, curriculum projects, state departments of education, Title III projects, and similar agencies.

The goals of the Ph.D training program are to produce professionals capable of (a) establishing and teaching evaluation and development programs and courses at the undergraduate and graduate levels in colleges and universities, (b) advising master's and doctoral students pursuing advanced degrees in

evaluation and development, (c) conducting basic and applied research on methodologies of evaluation and development, (d) consulting with educational agencies on these activities, (e) assuming leadership roles in professional organizations concerned with evaluation and development, and (f) maintaining an active interest in pursuing their own further education in the methodologies of evaluation and educational development. Such individuals would likely take positions within the academic world. Only half of the persons trained at the doctoral level would be trained towards such a goal. The other half would be trained towards the goal of working in such settings as school districts, state departments of education, regional laboratories, research and development centers, special projects, or private industry. The emphasis in their training would be on the administration and organization of educational evaluation and development, consultation on these activities, and acquisition of the skills necessary to perform evaluation and development successfully.

Broad goals of training at the master's level are to provide the non-supervisory manpower force for various agencies involved in evaluation and development. Persons trained at the master's level would not be expected to assume major responsibility for the activities of several other professionals engaged in evaluation and development. Emphasis in training at the master's level will be on imparting specific skills to trainees which can be coordinated in a team effort in evaluation or development.

Short-term intensive training institutes will aim at upgrading skills of persons already on the job or providing skills to persons seeking a redefinition of role in a school district or other

education-related agency. Emphasis will be given to training such persons to be perceptive and critical consumers of consultation and advice from evaluation and development specialists. It is expected that persons so trained will be able to carry out educational evaluation and development on a limited level with the assistance and counsel of persons trained at a more comprehensive and higher level.

Complexes of Skills

At a more specific level, one can identify groups of skills which will be imparted through the training programs in both evaluation and development; these are presented below.

Evaluation. In this section ten complexes of skills important to the successful pursuit of educational evaluation are identified.

1) Budgeting and managing human and material resources.

The typical educational evaluation is a large multi-faceted endeavor. In most instances it entails a financial budget exceeding that of many educational research projects. The activities of many technicians must be directed. It is essential that those who would engage in educational evaluation be trained in some of the techniques of project management and financing. This skill area will receive heavy emphasis in the doctoral program. At the master's training level, only those persons identified as potential directors of evaluation projects will receive similar training in this regard. For the bulk of trainees at the master's level and those participating in intensive institutes, this area will not be emphasized.

2) Identifying, at appropriate levels of generality, the goals of the program to be evaluated.

An important first step in any evaluation will typically be eliciting, from responsible parties, the goals and objectives toward which a program is directed. Identification of these goals must be more than a perfunctory attempt at soliciting verbal goal statements from program personnel. When done properly, it entails some of the most sophisticated technology of survey research and interviewing. At all three levels of training, emphasis will be given to the process by which goal statements can be identified in written documents, verbal interactions, and program plans.

3) Assessing the value of program goals.

In an evaluation, the goals of the program must not be accepted at face value, but must be regarded as elements of the program requiring direct evaluation. In some instances the justification of goals must come from empirical research in education or the social sciences. As an example, a program aiming at the inculcation of reading readiness skills in five-year-old children immediately raises the evaluative question of the facilitative effect on reading of the attainment of those skills identified as "reading readiness." The evaluator must know how to search for a justification of these goals in empirical research on reading instruction. He may find, for example, that the case for reading readiness has never been adequately established through empirical research. Hence, he may legitimately raise a question

of the justification of the program objective. In other instances, the evaluator may have to turn to nonempirical, nonbehavioral disciplines such as philosophy and law in seeking to evaluate program goals. A school system may set out as an objective to teach all children in the school the Christian ethic. The evaluator must be sensitive to, and should raise the issue (either privately with program personnel or publicly in his report), of the legal and philosophical issues concerned with the separation of church and state in the United States. Such goal evaluation requires broad education in the social sciences, philosophy, law, etc., and responsiveness to questions of value which are broader than those one can hope to investigate in the span of a single study.

4) Translating broad objectives into specific observable objectives.

General goal statements must be operationalized into specific statements of objectives. The onus of making this translation lies clearly with the evaluator, who possesses the technical skill for doing so, and not on program personnel, to whom the language of operationalization and behaviorism is foreign and unfamiliar. At all levels of training emphasis must be given to the technology of behavioral statements of instructional objectives. One may draw upon a large body of published literature in pursuing this goal and provide frequent practical experiences in the translation process.

5) Identifying standards or norms for judging worth.

The measurements and observations taken in an evaluation cannot be translated into judgments of worth without standards or norms. The formality of standards or norms may very greatly, but nevertheless a standard of worth is implied whenever judgment of worth is derived from an observation. Standards may be either internal or external to a particular evaluation. External standards are represented by collateral data with which observations and measurements are compared in deriving evaluations and judgments of worth. For example, a school system may desire racial balance in its schools and have attempted to achieve this balance through various means. The observation that 75 percent of all Negro pupils attend schools which have less than 25 percent of the student body Negro is an observation not yielding immediate evaluative meaning, since no standard for judging this degree of racial mixing exists. A standard for judgment could be found external to the evaluation. In data from Colemans's Equality of Educational Opportunity it might be found that in a representative sample of school districts across the nation, 95 percent of Negro pupils attend schools which are 90 percent Negro. By bringing external data to bear on the observations from an evaluation study, it can be seen that a much more satisfactory racial mix was achieved within the school district in question.

Training in the use of collateral data to derive external standards for judging value involves familiarization with data sources relevant to education. Trainees must become aware of various agencies and the data collection and reporting services

they render. They must also be trained in the critical skills necessary to evaluate properly credibility and reliability of external data.

In many instances standards for judging worth are to be found internally within an evaluation. The whole area of comparative experimental design is a means of establishing internal standards by which the worth of certain activities can be judged. One program is pitted against an alternative program, and the worth of the former is measured vis-a-vis the outcomes of the latter. The study of the principles of comparative experimental design then becomes relevant to the problem of establishing norms for judging worth in evaluation studies. Thus, it must receive appropriate emphasis in attempts to impart skills in this general area.

6) Monitoring programs to detect deviations from design or specified procedures.

It is important to know what one evaluates. It is insufficient to accept mere labels when one has invested large portions of time and money in the observation and judgment of outcomes. It is necessary that a program be monitored through site visitations, interview techniques, survey research methods, classroom observation schedules, etc., so that the evaluator is clearly aware of the degree to which a program being evaluated was made operational.

7) Selecting or developing and using valid techniques of measurement to yield information on outcomes.

The worth of an educational program lies in its outcomes. It is crucial that the proper outcomes be validly measured. Objective valid data on program performance is the sine qua non of any justifiable evaluation. The evaluator must have skill in selecting and evaluating those measurement techniques that will reveal objective data on outcomes. He must know when a measurement technique threatens to misrepresent a set of behaviors. Trainees must be provided with the critical skills necessary to judge the reliability, validity, and practicality of standardized and ad hoc measurement techniques. They must be able to construct and evaluate cognitive and affective measures for formative and summative purposes.

8) Employing appropriate techniques of data analysis.

The evaluator must be broadly knowledgeable in the area of statistical data analysis. He must have a clear understanding of the fundamentals of a variety of data-analytic techniques. He must be able to use standard packages of computer programs (such as the BIMD series), although programming skills per se are not essential.

9) Making recommendations as a result of the evaluation.

The evaluator's responsibility to evaluate does not end with the collection, analysis, and reporting of data.

The data do not speak for themselves. The perceptive evaluator acquires a valuable perspective on the educational program being evaluated through his long and intimate association with it. To fulfill his total evaluative responsibility, he must make the subtle and personal inferential leaps from those data he has gathered and those results he has observed to his personal recommendations for the conduct or continuance of a program. This is a skill not easily taught. The activity draws upon accumulated experience, wisdom, and judgment of the evaluator.

10) Writing the evaluation report.

Drafting the report of an evaluation for the relevant audience is an activity quite unlike the writing of a report of a research project. The evaluator is typically communicating to an audience which does not share his perspective, his grasp of technical topics, nor his interest in technical details. The responsibility to communicate findings rests more heavily with the evaluator than the researcher. The evaluator has to learn to adopt nontechnical language. He must refrain from over-reliance on tabular presentation of data analyses. He must avoid discursive commentary on test validity and reliability and other topics which his audience will not find central to their concerns. At this stage of the evaluation endeavor, the evaluator will play a role much more akin to the journalist than the scientist.

Development. The skills of the educational development specialist overlap at many points with those of the educational evaluator. However, there are important areas of difference. The following list of eleven skill areas is a delineation of many of the competencies required in a broad range of educational development projects.

1) Drawing on research results in planning developmental activities.

The developer will usually take his cues for particular developmental activities from published educational research. Research reports will never prescribe developmental activities; rather they inspire certain products and procedures. Developers must be sufficiently familiar and capable of reading the research literature that they can draw upon relevant findings in laying out the broad design of developmental activities. Research findings can never do more than suggest variables which may be manipulated in the production of educational products; but in order for the developer to make maximum use of research findings, he must be capable of interpreting the research literature. Hence, developers must be trained in the structure of the academic disciplines relevant to educational activities.

2) Conceptualizing systems, their elements, and interrelations among these elements.

Skills in special systems analysis are critical in the planning and implementation of educational developments.

Developers must be capable of seeing single developmental activities as parts of larger ongoing educational systems if the fruits of developmental activities are to be implemented and utilized to some good end.

3) Specifying desired performance outcomes of instruction.

The general concerns of administrators are the broad motivation for undertaking new educational developments. Seldom are these concerns stated in specific performance outcomes, before the educational developer is brought into the operation. Developers must be capable of translating the broad goals for an educational program, held by program personnel, into the specific observable outcomes which are to be expected from the educational product or program. Thus, training in the behavioral statement of objectives is crucial for developers at all levels.

4) Identifying alternative instructional media and techniques.

Educational development involves the manipulation of instruments to product desired outcomes. The media or instrumentalities available to educators must be carefully studied by developers. The range of alternative means to various ends can be fully employed. Developers must be well versed in traditional audio-visual media of instruction, organizational plans, and other instrumentalities through which desired ends of schooling are achieved.

5) Determining appropriate sequences of topics in instruction.

Where the complexity of instructional segments raises the problem of alternative ordering of elements to achieve desired goals, the developer must possess skills in logical task analysis and empirical confirmation techniques required to produce optimal orderings of topics.

6) Composing effective oral and written forms of instructional communications.

The heart of the developmental activity comprises the created acts through which the instrumentalities of instruction are used to bring about the desired objective. Developers must have the skills required to compose narrative or visual segments of instruction around which the teaching-learning process will center. Editorial skills, writing skills, skills in the creation of sequences of programmed instruction, etc., are the specific capabilities one can elaborate in this complex of skills.

7) Selecting or devising appropriate techniques for measuring outcomes.

It is usually necessary in developmental activities to devise new paper-and-pencil measures of learning to assess the specific objectives of instructional sequences. Developers require careful and intensive training in methods of test construction.

8) Designing and managing initial laboratory tests of developmental techniques and materials.

9) Designing and managing field tryouts and tests.

The new era of educational development differs from earlier, less successful attempts in that empirical data on product performance is systematically collected and fed back into repeated cycles in the development phase. The success of the endeavor often depends upon careful field tryouts in which data for revision are collected. Developers must possess skills in experimental design, test administration, and data analysis.

10) Reporting evaluation of outcomes from laboratory and field trials.

Skill must be acquired in identifying sources of data which bear on the critical development questions concerning which aspects of products or programs must be revised in order to meet performance standards. Once the relevant data are identified, great skill must be exercised in presenting data to development project personnel in comprehensible and usable form.

11) Interpreting evaluation findings.

The developer must be capable of specifying needed revisions, materials, and programs, from the study of field and laboratory trials. He must have a critical attitude toward evidence and be able to separate valid from invalid findings. Furthermore, he must be able to prescribe revisions in products from evidence of success and failure obtained in field and laboratory trials.

It is important to note that the field as a whole lacks traditional materials and arrangements for teaching many of these skills, ergo the critical importance of internships.

Specific Behavioral Objectives.

Coordinated with each complex of skills are sets of instructional activities. Each instructional activity breaks down further into several specific behavioral objectives. At this final level of specificity exist literally hundreds of statements of outcomes desired from instruction in the different types of training program. It is impossible in this limited space to enumerate all instructional objectives. Such objectives are in the process of being identified and ultimately will be identified for all areas of the training curriculum. As an illustration of objectives at this level, we present below the statement of behavioral objectives for one skill complex within the training program. These objectives are related to skills in comparative experimental design and data analysis. They are illustrative of the numerous instructional objectives which can be stated in all other instructional areas.

Behavioral objectives for instruction on the design and analysis of comparative experiments.

Upon completion of instruction the learner will be able to

1) Distinguish comparative experiments, associational studies, and status studies and describe the type of functional relationships that result from each.

2) Distinguish between internal and external validity and identify the threats to each as discussed by Campbell and Stanley.

3) Identify the experimental unit in an experiment and distinguish this unit from the unit of statistical analysis.

4) Determine the appropriate sample size needed to obtain a specified level of power in a given experiment and compute the power of tests previously conducted.

5) Explain how randomization and blocking function to minimize bias and how blocking further results in increased experimental precision.

6) Distinguish between main and nested classifications and between fixed, random, finite, and mixed models.

7) State the assumptions underlying a given analysis and describe the effects of failure to meet a given assumption.

8) Write the linear model and determine expected mean squares and degrees of freedom for a given experiment; compute sums of squares, mean squares, and appropriate F-ratios.

9) Describe the operations required to set up the following experimental designs:

a) randomized groups.
b) randomized blocks.
c) repeated measures designs.
d) Latin square designs.
e) balanced incomplete blocks.
f) analysis of covariance.

10) Graph and interpret two and three factor interactions.

11) Compute estimates of variance components for random classification analyses of variance.

12) Distinguish between planned and ex post facto comparisons between levels of a factor.

13) Select and compute the preferred multiple comparison procedure for planned and post hoc contrasts.

Training programs at the doctor's and master's level will have essentially the same broad goals and strive to achieve essentially the same complexes of evaluation and development skills. The two programs will differ in the depth to which

instruction is pursued on each of these objectives. Instruction at the doctoral level is intended to be broader and penetrating and will pursue each topic in greater detail. At the master's level some objectives will be pursued less deeply with less emphasis on theory. However, in the catalog of skills sought, there will be substantial overlap between training at the doctoral and master's levels in evaluation and development. Intensive training institutes will differ from the long-term master's and doctoral programs both in the coverage of topics and the depth to which they are investigated. Instructional objectives for institutes will be selected to meet the unique needs of trainees in these programs. An intensive training institute may emphasize only three or four skill complexes and not pursue these studies to the depth that they would be pursued at the master's and doctoral levels.

It is important to note that the Objectives Section above is somewhat traditional in its orientation toward broad goals, complexes of skills, and specific behavioral objectives. Not to be overlooked are important skills in broad humanistic areas, for example, skills in interpersonal relationships during evaluation and development, skills in task-oriented encounters, etc. These types of skill, by their very nature, are difficult to explicate and evaluate and have not been explicitly stated in this proposal. Nevertheless, the program staff is fully cognizant of the importance of this general area and will provide an instructional environment within which such humanistic skills will be nurtured and developed.

IV. Elements of the Training Program

The salient elements of the training programs designed are presented below. It can be noted that much of the training follows directly from the objectives stated in the previous section of this paper. (However, attainment of certain of the objectives will require internship experiences, team activities, and interaction between staff and students or among students.) Considered below are both the graduate degree training programs and the intensive training institutes.

A. Graduate Level Training in Evaluation and Development

Most of the formal academic work for all four degrees (masters in evaluation and development; doctorate in evaluation and development) will be conducted at the University of Colorado, Boulder campus. Each of the four programs will have unique elements, while elements of communality will also be in evidence. Several instructional techniques, for example, will be common to all four programs. They are enumerated below.

1) Classroom instruction: In each of the training programs, it is envisioned that approximately 40 percent or slightly less of the training experience will be conducted in formal classroom situations. The essential curricular experiences in the training of educational evaluators and developers are outlined later in this section. In many cases, new courses have been constructed, in order to be highly relevant for the program. Substantial amounts of knowledge and techniques from the social sciences have been incorporated into the courses. Additionally, trainees have options to elect independent study in areas of high interest to them.

2) Interdisciplinary orientation and seminars: It is anticipated that each doctoral trainee will be exposed to interdisciplinary work of at least two types: seminars or graduate level courses in other relevant disciplines; and advanced seminars in educational evaluation and development conducted by a team of scholars from other disciplines. Masters students will take part in a two-phase seminar in evaluation (or development).

The University of Colorado is fortunate to have an active program of interdisciplinary social-science research pursued by the Institute of Behavioral Science. In addition, research trainees in the graduate research training program of the Laboratory of Educational Research receive instruction in "general social systems" from Kenneth E. Boulding, one of the three founders of the General Systems Society, in the Department of Economics.

3) Seminars with peers: These problem-oriented noncredit seminars will be held weekly. Emphasis will be on using actual problem situations in evaluation or development, situations posed by practitioners attending the seminar seeking assistance. In many cases, trainees from all programs (including the current Title IV research program) will attend the same weekly seminar, thereby promoting facilitative communication, training, intense tutoring, and interaction among trainees in evaluation, development, and research.

4) Simulated materials, exercises, and problems in evaluation and development: Some materials are already prepared and will be used by trainees individually or in small groups. Other materials will be developed in the course of the project (see descriptions of Simulations I, II, and III in Section VII).

5) Internship/practicum experience: Such activities will be a key element in the training programs, amounting to approximately 50 percent of the trainee's time and effort. Internships will be offered in development and various evaluation roles. Receiving institutions will not provide any financial support for the intern, nor will receiving sites be paid by the project to house an intern. It is felt that the value of the work done by the intern for the agency should balance the costs of supervising and housing the intern. However, at certain development sites, actual instruction will be provided (Seminar in Educational Development); when the seminar is conducted by a development staff member, the necessary release time for him will be paid for by the project. Specific internship sites are outlined later in this section. In some cases, it will be desirable and possible to provide trainees with more than one practicum experience. In addition, all trainees will be exposed to many problem situations in the field and encouraged to formulate appropriate developmental (or evaluative) procedures.

The Masters Program in Evaluation

The master's degree program in educational evaluation is seen as a full year experience (two semesters and a summer session). Students selected will be of the highest possible quality (see Section VI). The bulk of the formal course work will be completed the first semester in residence, leaving the remaining six or seven months primarily for extensive field experiences as an intern as well as electives when feasible. The diagram below depicts the sequence of courses in the program.

Fall Semester	Spring Semester	Summer Session
1) Research Methods and Applied Statistics	5) Seminar in Educational Evaluation I	→ II
2) Educational Evaluation	6) Supervised Internship in Educational Evaluation I	→ II
3) Instrument Design and Development		
4) Educational Measurement		

The content and credit hours associated with each of the courses and experiences is designated below. (It might be noted that 1, 3, 5, and 6 are all new courses.)

1) Research Methods and Applied Statistics: 5 hours.

- Critical examination of current educational research literature.
- Role of empirical methods in education.
- Measures of central tendency, variability, and correlation.
- Inferential statistics.
- Experimental and quasi-experimental designs and analyses.
  * Controlling sources of internal and external invalidity.
  * Factorial designs and interactions.
  * Analyses of covariance.
- Survey research methods.
- Computer as a research tool (lab).
- Flow charting (PERT, etc.).

2) Educational Evaluation: 3 hours.

- Evaluation and research.
- Evaluation and measurement.
- The ethics of evaluation.
- Evaluation models.
- Simulated evaluation activities.
- Instruments in educational evaluation.
- Systems analysis.
- Cost-benefit analysis.
- Behavioral objectives.

3) Instrument Design and Development: 3 hours.

- Development and use of cognitive measures.
- Development and use of attitudinal measures.
  * Q-sort.
  * Semantic differential.
  * Likert scales.
- Development and use of sociometric measures.
- Conceptualization and use of unobtrusive measures.
- Reliability assessment methods.
- Validity determination.
- Item sampling.
- Criterion-referenced measurement.
- Self report measures.

4) Educational Measurement: 3 hours.

- Evaluation and measurement.
- Measurement in school settings.
- Teacher-made tests.
- Reliability.
- Validity.
- Types of items.
- Standardized tests.
  * Achievement.
  * General mental ability.
  * Special abilities (creativity, etc.).
  * Special aptitudes.
  * Interest and personality surveys.
- Informal measurement techniques.
- Projective tests.

5) Seminar in Educational Evaluation, I and II: Each 2 hours.

The focus of these seminars will be on applied problems in evaluation. Scheduled to coincide with the period that the student is an intern, it is expected that many presentations

will be by the trainees themselves on problems encountered at the internship site. Other sessions will provide exposure to persons filling roles in evaluation and development, roles which the trainee might select upon completion of his master's program.

6) Internship in Educational Evaluation, I and II: Each 2-6 hours.

It is obvious that the primary focus of the last six or seven months' training will be the internship, accounting for the bulk of the trainee's program during that time period. In some cases, a trainee can spend the full internship at one site and get experience working on various skills. On other cases, it will probably be wise to move students to a second site after several months at the first. Extensive supervision will be provided by the internship site staff (and sometimes instruction when it can be arranged); some supervision will also be conducted by the project staff.

A wide range of internship sites is available for evaluation students at the master's level. Among them are:

- John F. Kennedy Child Development Center, Denver, Colorado.
- Denver Public Schools, Denver, Colorado.
- Northern Colorado Educational Board of Cooperative Services, Boulder, Colorado.
- Colorado Department of Education, Denver, Colorado.
- Ford Foundation Field Office, Denver, Colorado.
- Office of Education Regional Office, Denver, Colorado
- Public school districts within fifty mile radius.

- Southwest Regional Laboratory for Educational Research and Development, Inglewood, California.

- Southwestern Cooperative Educational Laboratory, Albuquerque, New Mexico.

The Masters Program in Development

The master's degree program in educational development, also a twelve-month experience, follows the same pattern as the evaluation master's program. Most of the formal courses taken are the same; formative evaluation and instrument development are seen as major responsibilities of a master's level developer. The basic differences between the evaluation and development programs occur in seminars, electives, and internships. Again, most of the formal classes are completed in the first semester, as denoted in the diagram below. (All the courses are new except educational evaluation.)

Fall Semester	Spring Semester	Summer Session
1) Research Methods and Applied Statistics.	5) Seminar in Educational Development I ——→ II	
2) Educational Evaluation.	6) Supervised Internship in Development I ——→ II	
3) Instrument Design and Development.		
4) Educational Development		

The course descriptions for 1, 2, and 3 are given in the evaluation section and are not repeated here. Specifics on the other courses follow.

4) Educational Development: 3 hours.

- Development and research.
- Development and evaluation.
- Development and diffusion.
- Formative evaluation and development.

- Behavioral objectives.
- Interpretation of research in planning development.
- Role of media in development.
- Design of laboratory and field tryouts.
- Report and interpretation of findings.

5) Seminar in Educational Development, I and II: Each 2 hours.

Applied problems in product or technique development will serve as the focus of these seminars. Running concurrently with the internship experience, the seminar should represent an excellent vehicle to carry internship problems to the seminar table. Also, efforts will be made to expose trainees to development role incumbents. This seminar will be facilitated by an individual, or team, actively engaged in the development process. Focus in the seminar will be on creative solutions to development problems, zetetics, and the like.

6) Internship in Educational Development, I and II: Each 2-6 hours.

During the extensive block of time established for the internship experience, it is desired that trainees be exposed to many considerations in development and that they become acutely and deeply involved in the process of development. The intention is to leave many trainees at a single site for the entire six or seven month period (except in unusual cases) so as to increase the probability of involvement in an on-going program. Supervision will be provided by the intern-site staff, and supervision will also be forthcoming from the project staff.

Considering that major curriculum development does not take place in many agencies, the existence of several national curriculum projects in the Boulder area enhances the types of internships that can be offered. Consider the list of sites below:

- Biological Sciences Curriculum Study, Boulder, Colorado.
- Earth Sciences Curriculum Project, Boulder, Colorado.
- Social Sciences Education Consortium, Boulder, Colorado.
- John F. Kennedy Child Development Center, Denver, Colorado.
- Denver Public Schools, Denver, Colorado.
- Northern Colorado Educational Board of Cooperative Services, Boulder, Colorado.
- Colorado Department of Education, Denver, Colorado.
- Public school districts within fifty mile radius.

The Doctorate in Evaluation

The Ph.D. program in educational evaluation described here, and the Ph.D. program in development described below, are designed to meet all requirements of the Ph.D. degree established by the University of Colorado Graduate School and the School of Education. Language and other examination requirements will be met by all candidates for these programs. It can be noted that the relative weight on internship training for credit is considerably greater than usual, and the locus of the internship is often in nonacademic settings.

Each component of the program is described in detail below. It is assumed that persons entering the program will have either the master's degree in evaluation or its equivalent. If they do not, they will have to take the master's degree too, resulting in at least a three-year program.

All doctoral students in evaluation will be expected to complete the following courses, or equivalents (in addition to the Master's level courses).

1) Intermediate Statistical Methods: 3 hours.

- Sampling theory and inferential statistics.
- Advanced applications for the testing of hypotheses regarding central tendency, variability, proportion, correlation, and normality.
- Chi-square, and other nonparametric methods for independent and related observations.
- Special methods of correlation.
- Multiple regression and prediction.
- Introduction to the analysis of variance.
- Computer programs for statistical analysis.

2) Intermediate Educational Measurement: 3 hours.

- Reliability and validity theory.
- Empirical estimation of reliability and validity.
- Interpretation of standardized tests.
- Norms.
- Special problems in assessing intelligence, achievement, interest and personality.

3) Survey Research in Education: 3 hours.

- Non-experimental techniques of behavioral and social sciences with applications in educational research and evaluation.
- Latest developments in survey research methodology.
- Inferring causation from survey data.
- Sociological sampling theory.
- Maximizing response rates.
- Ex post facto techniques.

4) Experimental Design and Analysis: 3 hours.

- Experimental and quasi-experimental designs in educational research.
- Selecting an appropriate statistical test.

- Power and power efficiency.
- Randomization and control.
- Randomized groups and blocks designs.
- Multiple comparisons.
- Factorial experiments and interaction.
- Repeated measures designs.
- Analysis of covariance.
- Effects of assumption violations.
- Related computer programs for statistical analysis.

5) Advanced Seminar in Educational Evaluation: 2 hours.
- Selected topics for advanced study in educational evaluation.
- Evaluator-client relationships.
- Role of the evaluator.
- Identification of audiences and influence groups.
- Reporting evaluative information.

6) Seminar in Assessment Research: Variable credit (Sociology course).

The seminar will be concerned with methods of assessing the effectiveness of action programs conducted in various institutional sectors of the community. Basic principles of research design, measurement, and administration in the behavioral sciences will be applied to the situations likely to be encountered when social research is conducted in an action setting. Extensive case material will be utilized.

To allow maximum flexibility in tailoring individual programs to meet needs of individual students, several electives, determined by the student and his advisor, can be taken. These might include more internship time, electives, independent study, or seminars. Suggested electives of the latter type are listed below, although others may be included (of special value would be courses dealing with investigative methods in other disciplines such as economics, anthropology, sociology, etc).

1) Human Learning: 2 hours.

Theory and research in human learning with implications for educational research and classroom learning.

2) Advanced Psychological Foundations of Education: 2 hours.

On the learner, the learning process, and learning theory as it applies to the educative process.

3) Advanced Social Foundations of Education: 2 hours.

An evaluation of the social values and forces in American society that shape or influence the aims, philosophies, methods, content, issues, and problems of the American educational enterprise.

4) Curriculum Construction: 2 hours.

Methods of formulating curriculum programs. Procedures for locating, organizing, and summarizing data concerning social, economic, and personal problems related to the future needs and interests of youth and adults.

5) Seminar in Curriculum Construction: 2 hours.

(See description under development doctorate).

6) Advanced Seminar in Educational Development: 2 hours.

(See description under development doctorate).

7) General Social Dynamics: 3 hours (Economics course).

The development of a general theory of the dynamic processes of the total social system.

8) Computer Decision Modeling: 3 hours (Computer science course).

Application of the methods of computer science to problems in industrial management. Emphasis is placed on simulation as

a method for studying the behavior of dynamic systems and the use of optimization models for their control.

9) Computer Oriented Decision Modeling: 3 hours (Management science course).

Application of the methods of computer science to problems in management. Emphasis is placed on simulation as a method for studying the behavior of dynamic systems and the use of optimization models for their control.

10) Seminar in Management Science: 3 hours (Management science course).

Application of operations research methods.

11) Personnel Management - Policy and Practice: 3 hours (Manpower management course).

Development of policy and procedures for various personnel management functions. Application of control and evaluation techniques to such areas as organizational structure, recruitment, placement, selection, training, wage administration, performance rating, morale, and management development. Quantitative analysis, cases, and simulation techniques are employed.

12) Management of Personnel Systems: 3 hours. (Manpower management course).

Analysis, design, and quantitative evaluation of various organizational structures and systems involved in such personnel functions as recruitment, selection, training, wage administration, performance evaluation, communications and incentives.

13) Attitude Assessment I and II: 2 hours each (Psychology courses).

A review of methods for measuring attitudes and their psychometric bases.

14) Multivariate Analysis: 3 hours (Psychology course).

Analysis of experimental and observational multivariate data with emphasis on correlational approaches and the use of multivariate procedures in the design of research.

15) Sampling and Inference: 3 hours (Statistics course).

Stresses methods of sampling used to make reliable inferences. Sampling distributions, sample surveys, sample design, and related inference techniques.

16) Statistics of Social Relations: 3 hours (Sociology course).

Statistical and probabilistic methods relevant to social processes; time series analysis and stochastic processes; methods relevant to social structure; random nets; path analysis; and segregation measures. Topics selected for their particular utility in handling sociological problems.

17) Seminar in Sociological Analysis of Organizations: Variable credit (Sociology course).

Analysis of theoretical and research issues in the study of organizations of all types. Systematic review of literature.

Although the various courses provide necessary background, the heart of training educational evaluators lies in the actual planning and conduct of evaluations as part of a continuous, supervised internship in evaluation. Each student will be expected to enroll in two internship sections, one per semester, during an entire calendar year:

this amounts to approximately twenty hours per week of supervised internship time over the twelve-month period.

Internship sites will include all consortium units; particularly relevant in this case will be the regional labs and CIRCE. Under joint supervision of internship instructors and senior staff members of the consortium unit, interns will participate in all phases of planning, conducting, and reporting evaluations. The internship is designed (a) to provide a context within which to integrate the knowledge gained through coursework and seminars and (b) to provide an opportunity to learn under supervision, evaluation techniques, skills and sensitivities that can only be learned through actual evaluation experience. Specifically, interns will be provided with opportunities to: (a) interact with clients to determine their evaluation needs, (b) identify (at appropriate levels of generality) the goals of the program to be evaluated, (c) help program staff assess the value of these goals, (d) help program staff translate goals into specific, observable objectives, (e) determine information needs of program administrators and alternatives in relevant decision situations, (f) identify sources of relevant information, (g) design data-collection techniques for use in collecting needed information, (h) identify standards or norms for judging worth of information, (i) developing and administering evaluation instruments, (j) monitoring on-going programs to detect deviations from design or specified procedures, (k) employing appropriate techniques of data analysis, (l) making recommendations based on evaluation data, and (m) reporting evaluation results to appropriate audiences. Benefits of such an approach accrue

not only to the intern but also to consortium units, who are provided with services of highly talented apprentices.

Further, it is anticipated that students will find that they need varying types of information and help either not provided through formal courses and seminars or not provided in time to meet pressing needs related to their internship work. Positive feedback from students will be elicited to assess immediate needs of students not met through other aspects of the program. If needs are unique to one or two students, they will be routinely provided by faculty internship supervisors. However, whenever needs are found to be common to several students, informal seminars will be organized on an ad hoc basis, with duration determined by the topic. For example, needs for assistance in managing complex evaluation projects may lead to a week-end seminar on applying PERT techniques to manage human and material resources. All such seminars will be conducted by University of Colorado faculty, guest instructors from consortium units, and outside instructors as necessary.

In addition to such ad hoc seminars organized around specific needs, an informal noncredit seminar will be held on a weekly basis for all students, faculty, and consortium members able to attend. The agenda will be determined by the students and may include substantive content or organizational (business) details.

Finally, each student will be expec[illegible]o complete an original dissertation on some aspect of educatio[illegible]valuation.

The Doctorate in Development

At this point in time, the specifications for the doctoral program in development are not as exact and prescribed as they are for either of the master degree programs or for the doctorate in evaluation. However, it is felt, since there will be no doctoral students during the first operational year, that it will be possible to develop a quality program at this level. It can also be noted that since there will be six master's candidates to select from, the student selected undoubtedly will be outstanding and will assist materially in development of a sound program. In addition, the experience gained during the initial year will provide useful input for the ingredients needed in the doctoral program. It is assumed that persons entering the doctoral program will have completed the master's program in development, or its equivalent. If this is not the case, those courses will have to be taken.

If the candidate is very certain that he is entering development in a specific subject area, he might wish to take additional courses in that particular discipline that will better prepare him. Generally, however, doctoral students in development will be expected to complete the following courses, or equivalents (in addition to the Master's level courses):

1) Intermediate Statistical Methods: 3 hours.
   (See description under evaluation doctorate).
2) Intermediate Educational Measurement: 3 hours.
   (See description under evaluation doctorate).
3) Advanced Seminar in Educational Development: 2 hours.
   - Creative solutions to existing problems in development.

- Selected topics for advanced study in educational development.
- Establishment and maintenance of human relationships.
- Selection of appropriate instructional and media techniques for developing educational products and/or processes.
- Direction of the work of production personnel.
- Management of laboratory and field tryouts.
- Interpretation and description of products for different audiences.

4) Seminar in Educational Media: 2 hours.

- Definition of problems in educational media.
- Solution of problems in educational media.
- Experimental use of media.
- Research on use of media.

5) Curriculum Construction: 2 hours.

(See description under evaluation doctorate).

6) Seminar in Curriculum: 2 hours.

- Advanced theory in curriculum construction.
- Advanced practice in curriculum construction.

Additionally, there is a wide range of electives (or independent study) available for the doctoral student in development; for the most part, they are the same as those for the doctoral student in evaluation. Several possible electives are listed below; the reader can refer to the evaluation section for more complete course descriptions.

1) Survey Research in Education: 3 hours.

2) Experimental Design and Analysis: 3 hours.

3) Advanced Seminar in Educational Evaluation: 2 hours.

4) Human Learning: 2 hours.

5) Advanced Psychological Foundations of Education: 2 hours.

6) Advanced Social Foundations of Education: 2 hours.

7) General Social Dynamics: 3 hours (Economics course).

8) Computer Decision Modeling: 3 hours (Computer science course).

9) Computer Oriented Decision Modeling: 3 hours (Management science course).

10) Seminar in Management Science: 3 hours (Management science course).

11) Personnel Management - Policy and Practice: 3 hours (Manpower management course).

12) Management of Personnel Systems: 3 hours (Manpower management course).

13) Sampling and Inference: 3 hours (Statistics course).

14) Seminar in Sociological Analysis of Organizations: Variable credit (Sociology course).

15) Seminar in Assessment Research: Variable credit (Sociology course).

16) Attitude Assessment I and II: 2 hours (Psychology courses).

17) Multivariate Analysis: 3 hours (Psychology course).

It is important to note that there are at least six strong internship possibilities for the doctoral student. Unlike the internship at the master degree level, it is feasible and probably desirable that the doctoral candidate take more responsibility in directing and managing certain aspects of the actual development process when at the internship site. Thus, he might select any one of the following six sites (or any other that seemed suitable):

- Biological Sciences Curriculum Study, Boulder, Colorado.
- Earth Sciences Educational Program, Boulder, Colorado.
- Social Sciences Education Consortium, Boulder, Colorado.

- Southwestern Cooperative Educational Laboratory, Alburquerque, New Mexico.

- Southwest Regional Laboratory for Educational Research and Development, Inglewood, California.

- Center for Instructional Research and Curriculum Evaluation, University of Illinois, Urbana, Illinois.

In meeting the thesis requirement at the doctoral level, the student in development quite likely would conduct research bearing on the methodology of development. The extreme paucity of published research on development makes it critical that the doctoral candidate do original research in this area before he completes his advanced degree.

Number of trainees in the graduate level programs

It is apparent that a single training program cannot hope to play the numbers game and come out ahead in the long run in terms of the quality of the students thus trained. In attempting to determine feasible and compatible numbers of students to involve in this training program, the real constraints seem to hinge primarily on the quality of the trainee rather than on budgetary considerations. Additionally, it is desired to train large numbers of persons at the masters level and return them to the field so that they may actively commence to have an impact in the roles to which they are assigned. Establishing a balance, then, between the masters and doctoral program is also a very real consiera-tion.

The table on the next page indicates our current thinking as to the desired number of trainees. As can be noted in the table, it is proposed that eighteen students be supported each year. No doctoral students would be involved during the first year; rather from the twelve masters

Table 4 Numbers of trainees by year, field, and degree level.

Field	Degree Level	Year One	Year Two	Year Three	Year Four
Evaluation	Masters	12	10	8	8
	First year doctorate	0	2	2	2
	Second year doctorate	0	0	2	2
Development	Masters	6	5	4	4
	First year doctorate	0	1	1	1
	Second year doctorate	0	0	1	1
Cumulative Number of Graduates in Evaluation and Development Entering the Work Force	M.A.	15	27	36	45
	Ph.D.	0	0	3	6

students in evaluation two would be identified to continue on to the doctorate while one of the six masters students in development would be so identified. It would take three years until the first doctoral students complete the program. In the meantime, however, considerable number of students at the masters level will be entering the job market (ten the first year, eight the second year, and six each subsequent year in evaluation; the corresponding figures for development would be five, four, and three). This is making the assumption that the doctoral students are only selected from successful master candidates at Colorado. This is not likely to always be the case; thus in some years the number of master students entering the job market would be greater than those indicated above. The reader should note that the figures in the table are meant to be suggestive only. It is anticipated that certain years it might make more sense to train more evaluators or fewer, or more developers or fewer depending on the apparent job market and also on the interest and qualifications of students entering the program. However, it is suggested that the total number of students in the program at one time be limited to eighteen. Past experience would suggest that other students without stipend support will become interested in the program and probably will participate. There is clearly some point at which the group becomes too large for each of the individual trainees to feel that they are undergoing specialized high-quality training.

B. Intensive Training Institutes

The general nature of the intensive training institutes proposed revolves about the word flexibility. Considerable flexibility is seen in terms of the content of the institutes, their length, and the particular time of year in which they are scheduled. Most important of these is content flexibility.

The intention is to achieve content flexibility by designing a series of training package modules or mods. These mods will typically take from one day to three or four days to complete. Later in this technical paper, more information on the development on these training packages is given (Section VII). For our purposes here we need only note that, for each specific group coming to an institute, it would be possible to select and fit together those mods most appropriate for the group to be trained. Thus, it is likely that no two workshops will be identical in composition. To the extent that time permits, it will be desirable to have representatives from the potential training population meet with project staff several months before the proposed training, in order to examine and select those mods most relevant and appropriate. It is felt that this flexibility in terms of content, and the ability to match the content to the specific needs of the trainee group, will prove extremely advantageous.

Flexibility is also envisioned in terms of the length of the institute and also the time period in which it is scheduled. It is felt that two weeks will be a likely average length for institutes, but there is no reason why certain institutes might not be scheduled for only a few days and others for three or four weeks. It is also conceivable that

a group might attend an institute for two weeks and return several months later for an additional week or so. It should be noted however, that although the lengths of the institutes might vary, it is intended that the intensity will not and that each training program will be conducted at an intense level. The time period in which the institute is scheduled is also flexible, subject to the constraint that there are certain times of the year when most educators cannot attend because of peak work responsibilities. Thus, institutes will be scheduled to avoid such barriers.

For the first twelve to eighteen months, training in the institutes will be a secondary consideration with development being primary. Those workshops that are conducted in the first year will be primarily to perfect the training model and further refine the materials being developed, rather than merely to train personnel. Even though training is considered secondary during the first year, it is expected that it will be of tremendous benefit to those involved. It is probable that a two-week workshop will be held in June, 1971, and a second two-week workshop in August, 1971. Emphasis will be on development of the workshop model, procedures, and materials, rather than on the training alone.

It is believed now that at some point it will be possible to run institutes in tandem thereby effecting considerable savings in terms of staff resources utilized to conduct the training. Thus, two or even three workshops might be held simultaneously at the same site. When the groups involved have selected the same training mod, then that training mod could be presented to the combined groups rather than to each group separately. This being the case, it would be possible to

schedule institutes during the "off season" and yet still reach a large number of trainees. It is intended that each workshop contain about twenty-five to thirty participants, on the average, although trainee groups smaller or larger than this on occasion can be expected. Commencing with the second year, it is expected that about 200 trainees might receive two or more weeks of instruction annually.

It is important to note that given the flexible content approach and the training package modules, it is somewhat academic to specify whether an institute is being held in development or in evaluation. It is probably true that most persons filling evaluation roles have some development responsibilities and vice-versa. Thus, in selecting the mods most appropriate for itself a group is liable to select from both development and evaluation topics.

Several possible models of follow-up will be examined to determine which are most effective in an absolute sense and also which are most effective from a cost efficiency standpoint (see Section VIII). Presently, it is conceived that four alternative models be compared experimentally, with a portion of trainees randomly assigned to each. Treatment One might be considered the traditional treatment, that is, essentially no follow-up at all. Treatment Two might consist of letters, and phone contacts to participants. Treatment Three would include Treatment Two plus having participants return to training settings for short (possibly one day) follow-up sessions. Treatment Four would consist of Treatment Three plus assistance and supervision provided to the trainee at his home site. It is apparent that these four models represent rather dramatic differences in terms of staff resource commitment

(from essentially zero commitment in Treatment One to extensive commitment in Treatment Four).

Recall also that on occasion institutes might be held for special groups. For example, it would not be particularly difficult after a particular instructional mod had been tried out and modified to present just that mod to a group of early childhood fellows attending the course at the John F. Kennedy Child Development Center. Also, short-term institutes might be planned of an evangelical nature, say for supervisors of evaluators or for school decision-makers (recall Assumption 6 in Section II that stressed the need for establishing a climate supportive of vigorous evaluation and development activities). For example, a two or three day institute might be set up early in the Spring for a group of assistant superintendents in charge of instruction, persons who have key decision making roles in the school hierarchy. The purpose of the session would be to acquaint the group with the types of demands they could place on an evaluator and the assistance that an evaluator could provide to a decision maker. The session would also be used to create enthusiasm for a subsequent institute that would be conducted in the Summer and to which evaluators from the same school district (as the decision makers) would come. It is likely that continuing efforts will have to be undertaken to facilitate the proper use of evaluation and development trainees when they return to their normal role settings. Note, too, that the two short institutes used as examples above (for Kennedy fellows and assistant superintendents) would not require stipend payments to trainees. This type of "free" institute can be used to

extend the impact of the evaluation / development message without expending massive funds.

The manpower to conduct the intensive training institutes will come from three sources. The first source will be the Project Director and the Associate Director for Training Institutes, each of whom will take some teaching responsibilities. The second source will consist of persons that might be called the consultative staff. In most cases, these persons will include the same professionals that develop the training materials modules. That is, if Professor A develops the mod on survey research, it is intended that he would also conduct the training, at least initially, in that mod. It is expected that there will be several sources of consultative staff. To the extent that their time commitments allow, the other two Associate Project Directors (Glass and Worthen) and the Project Formative Evaluator will be involved as instructors; it is expected that key personnel from CIRCE might be involved; consultative staff in development will be sought from the national curriculum projects; etc. It is thought that because of the great flexibility in scheduling these institutes and because of their sporadic nature, it would be best to obtain instructors on a consulting basis rather than to hire permanent instructional staff who necessarily would be idle during large portions of the year.

The final source of manpower for the short-term training institutes would be the graduate students in the program. It is felt that the doctoral level students can assist materially in the instruction of the trainees and in the process obtain excellent training for future career activities. The master students on the other hand, could serve

ideally as tutors for trainees that are having particular difficulties with material being presented. The intensive training institutes should function to involve graduate students intimately with the process of training in evaluation and development; this involvement is seen as very desirable.

Finally, it should be noted that many of the materials development activities (Section VII.) will have direct relationships to the institutes. The simulation materials developed might function either at an institute or be sent to and worked on by trainees at their home site. The training package mods are being developed primarily for the institutes although it is expected that they will be useful in other settings once developed. It is felt at this time that as many as sixty days of instruction might be developed. Likely topics for mod development and probable length in days is indicated in the list below:

Training Package Module	Probable length (in days)
1. Identification of program goals.	1.0
2. Design of evaluations.	2.5
3. Design of laboratory and field tryouts.	1.5
4. Flow charting and networking.	2.0
5. Formative evaluation techniques.	2.0
6. Development and use of behavioral objectives.	2.0
7. Program monitoring.	1.5
8. Observational techniques.	2.0
9. Standard measuring devices.	2.0
10. Instrument development.	4.0
11. Survey research.	2.5

12.	Matrix sampling.	2.0
13.	Measures of central tendency, variability, and correlation.	2.5
14.	Factorial designs and interaction.	2.5
15.	PPBS.	2.0
16.	Cost-benefit analysis.	2.0
17.	Role of media in development.	2.5
18.	Composition of effective instructional communications.	1.5
19.	Data analysis (computer as a research tool).	3.0
20.	Interpretation of evaluation results.	1.5
21.	Interpretation of development results.	1.5
22.	Human relations in development and evaluation.	2.0
23.	Evaluation from a decision-makers standpoint.	2.0
24.	Open	12.0
		60.0

V. Staff of the Training Program

It is envisioned that six staff members at the University of Colorado will make up the basic continuing staff component for the training program. It can be noted that the full-time-equivalent of the six positions is 3.3. Some might argue that three full-time persons (instead of the two full-time and four part-time staff envisioned) would comprise a more logical staff structure. However, the design staff opted for this alternative because of the diverse nature of the roles assigned to the staff, roles unlikely to be easily merged, and because the resultant staff structure seemed more flexible than the three full-time staff model. The six positions are listed below as well as time commitments and responsibilities with regard to the training.

1) Project Director. Dr. William L. Goodwin will direct the project (see vita in Appendix A). This is contemplated as a 70 percent commitment. Responsibilities of the Director include directing the training program, both graduate and institute, supervising the materials development, recruiting trainees (see Section VI below), and teaching one evaluation course. The Project Director will also be the chief administrative officer for the advisory board to the consortium (see Section IX below).

2) Assistant Project Director. It will be necessary to recruit someone, probably a recent Ph.D., to fill this full-time position. In addition to being familiar with all aspects of the program so he can respond for the Director in the event of the Director's absence, the Assistant Director will

assist in supervising the materials development activities and will teach two courses in the graduate training program.

3) Associate Director for Intensive Training Institutes. It will be necessary to recruit someone to fill this position, probably a recent doctorate in the field of evaluation, development or teacher training. This is viewed as a 100 percent time commitment. The responsibilities of the Associate Director for Intensive Training Institutes are primarily concerned with all elements of the institutes including recruitment, administrative arrangements for trainees, liaison with organizations providing the trainees, liaison with institute instructors, etc. In addition, he will teach one graduate course. One possibility for this position is Dr. Ronald Anderson of the University of Colorado. A commitment on his part is contingent upon many variables, so a vita is not included for him at this time.

4) Associate Director for Graduate Instruction. This position will be filled by Dr. Gene V Glass, and is contemplated as a 20 percent time commitment. Although this time commitment may seem small it should be recognized that Dr. Glass will be involved in the program more than that, through his teaching of several of the key courses; however, the University will be paying for the teaching of these courses, as they also will be offered to other than training program students. The Associate Director for Graduate Instruction has the responsibility of working with the professors who are conducting the graduate instruction to insure continuity in the program. Additionally, Dr. Glass will function in an advisory capacity in terms

of the training mods that are to be developed; correspondences between the graduate instruction and certain of the mods should be considerable. Dr. Glass' vita is given in Appendix A.

5) Associate Director for Graduate Student Liaison and Program Planning. This position is listed as one fifth of a full-time equivalent and will be filled by Dr. Blaine R. Worthen. As was the case with Dr. Glass, Dr. Worthen will be involved to a greater extent than the fifth-time infers, for courses that he will teach will be paid for by the university. The key responsibility of this Associate Director will be to maintain frequent and facilitating contact with all the graduate trainees in the program; this includes assisting them to plan their internship experiences. It is expected that Dr. Worthen will function as a resource person particularly sensitive to the problems of the graduate trainees in evaluation and development and to the needs of the various internship sites. He will also provide information directly to the Associate Director for Graduate Instruction relative to the student opinions of instruction and the student ideas upon improvement of instruction. Dr. Worthen's vita is given in Appendix A.

6) Program Formative Evaluator. The position of Program Formative Evaluator will be filled by Dr. Kenneth D. Hopkins and is viewed as a 20 percent time commitment. Dr. Hopkins vita is attached with this technical paper in Appendix A. There appears to be a critical need to provide, within the project structure,

one person who's chief concern is formative evaluation of the ongoing training program, both at the graduate level and the intensive training institutes. Functioning rather independently as a data collector, synthesizer, and analyzer, the Formative Evaluator has the responsibility of inputing his interpretations to the Project Director on how the programs might be improved, which are working well, which are working poorly, etc. It is thought that the program will benefit substantially from inclusion of this "built-in critic." Feedback from many sources will flow through the Program Formative Evaluator back into the system to be acted upon. To the extent that Dr. Hopkins needs assistance in carrying out certain formative evaluative studies, it is expected that graduate students in the program will have the requisite technical skills to assist him.

By examining the vitae in Appendix A, the extensiveness of the accumulated experiences of the designated staff and the variance in their backgrounds is apparent. It should also be noted that Drs. Glass, Worthen, and Hopkins have had extensive experience via the ESEA Title IV training program at the University of Colorado's Laboratory of Educational Research, while Dr. Goodwin spent three years as one of several departmental staff serving on master's level Title IV research training program at Bucknell University. Among other significant evaluation activities, the staff of the Laboratory of Educational Research, under the direction of Dr. Gene V Glass, recently completed the analysis of the data and the evaluation report on Title I of the ESEA for 1968-69,

under contract with the Bureau of Elementary and Secondary Education, U. S. Office of Education. Dr. Glass also served as AERA presessions chairman in 1968. Notable also are the relevant professional experiences of the staff: Dr. Blaine R. Worthen's tenure as Associate Director of the Evaluation Center at the Ohio State University;Dr. Worthern's work with Dr. David Clark in writing Preparing Research Personnel for Education; Dr. Worthen's directorship of the AERA Task Force on Research Training; and Dr. Hopkins' lead role in the 1968 AERA design presession. Further, Drs. Glass, Worthen, and Hopkins have conducted several short-term training programs for preparing evaluators under the Educational Professions Development Act and other authorizations. Dr. William L. Goodwin has directed four large scale projects (the SESAME projects) that had heavy developmental components; in addition he has served as an advisor to the State of Colorado in evaluating performance contracting being carried on in several school districts. Note also that Drs. Glass and Worthen have been actively assisting the Colorado State Department of Education over the past two years in planning and conducting an unusual statewide assessment program. In total, it is felt that proposed staff for the graduate level training has a unique assortment of abilities and experiences that complement one another.

It can be noted in the budget that there also is the intention to employ as an instructor for graduate trainees someone from one of the national curriculum projects to conduct seminars in educational development. This might prove to be a team operation with more than one person involved. The critical thing to note is that these key development courses will be taught by professionals with real expertise

who are actually engaged in extensive program and curriculum development.

Turning to the question of the staff for the intensive training institutes, it can be noted that some of the teaching would be done by the Project Director and by the Associate Director for the Institutes. It is conceivable that Drs. Glass, Worthen, and Hopkins will have additional time to commit to such activity, outside of the time that is already committed to the project. Supplemental staff clearly will be needed.

In effect, instructors will be sought for the intensive institutes who have two competencies: instructional abilities and skills in the development of materials. Thus, whoever develops the training package mod in the development of behavioral objectives would also be expected to conduct the institute training in that area. Possible sources of instructors and materials development include the national curriculum projects, the John F. Kennedy Child Development Center, State Departments of Education, Regional Laboratoties and Centers, and, notably, CIRCE. It is fully expected that it will be possible to assemble a powerful staff for each of the short-term institutes. From the standpoint of cost effectiveness, it probably will not be possible to have the institute instructors available during the entire course of a two to four week institute. Rather, it is hoped that the continuity that is so important for the trainees, can be provided by the Project Director and the Associate Director for the Intensive Institutes.

It can be noted that with the planned flexibility in terms of when institutes are held, it is quite likely that high caliber professionals can be attracted to do the materials development and also to conduct the

instruction. The revised and liberalized copyright arrangements now possible with the U. S. Office of Education should make the role of developer/instructor even more acceptable to many persons.

Finally, it is of special importance that much instructional assistance will be possible because of the high quality of the fellows that will be on this training program and also those fellows on campus undergoing research training. Masters students might serve well as tutors and instructional assistants in connection with the institutes. Doctoral trainees likely will take instructional roles, for both institute participants and master's level students. This valuable cadre of bright manpower should not be overlooked.

Let us now consider the recruitment and selection of the trainees who will work with the staff cited above.

VI. Recruitment and Selection of Trainees

Considered below are plans for the recruitment and selection of trainees, both for the graduate program and also for the institutes. It can be noted that the procedures for the graduate programs are stringent and quality-oriented. On the other hand, it is felt that for the intensive institutes, emphasis should be placed primarily upon the extent to which the prospective trainee either (1) has evaluation and development responsibilities as part of his current job description or (2) will be assigned evaluation and development responsibilities as part of a new or redefined role.

Recruitment and Selection of Graduate Students

Recruitment methods to attract graduate students to the program will include the distribution of brochures to leaders in the field of education who might have contact with prospective fellows, to state departments of education, and to major school districts. It is expected, however, that referrals from persons in other institutions that have become familiar with the program will become the single most important means of recruiting good prospects. Another source that has proved valid in the past is the recommendations of present fellowship students, already in the program. In some respects, it is believed that the mailing of literature is necessary to initiate the program, but that "flashy" brochures become less important with the passage of time. It is hoped that the methods of dissemination of information about the program are effective enough to attract applicants from a diverse group of potential fellows. It is also intended that announcements of the program will be placed in the AERA Educational Researcher, the NCME Newsletter, the Phi Deltan Kappan, and the APA Monitor.

It is important to recruit nationally for this program. Diversity of backgrounds and of experience brought to the program is a strength that often is not possible with only local recruiting. Likewise, the belief is that the higher the quality of trainee input, the higher the quality of the total program is likely to be. Applications for the fall of 1971 should be received by March 15, 1971.

Selection of fellows will be a complicated task, and each applicant will be considered individually. Although several entry characteristics desired for fellows are given below, it should be stressed that any, or all, of them might be waived in the case of a particular individual.

Desired is high academic potential as reflected in a minimum average grade point of 3.0 (on a 4.0 scale), with at least a 3.5 grade point average on any graduate work completed. (Note that persons with any undergraduate major will be eligible to apply.) Additionally, it is felt that a minimum combined score of 1125 on the quantitative and verbal aptitude sections (with a quantitative score of not less than 560) on the Graduate Record Examination would be desirable. This standard might be higher if the intent was only to train doctoral students. However, with the intention to return the predominance of the trainees to the field after the master's degree, it is felt that a combined score of 1125 is appropriate.

A second criterion governing the selection of fellows will be the nature of the background experience of each of them. Looked upon favorably will be strong interdisciplinary preparations. Additionally, academic work in the behavioral or social sciences will be regarded positively. For evaluation trainees, some college work in mathematics

is very desirable unless the GRE quantitative score is indicative of excellent aptitude in the quantitative area. It will not be necessary that the persons applying have had extensive experience in education; it is likely that previous experience in education will not be a critical factor in determining fellowship recipients.

The final criterion used in selecting fellows will be judgments made on their apparent interest and commitment to a career in educational evaluation or development. To this end, a candidate's professional objectives will be examined as will any supporting letters of recommendation that are submitted. It might be noted that other applicant qualifications considered will be his age, the clarity with which he expresses himself in his statement of objectives, and how he conducts himself in an interview (if such an interview with the Project Director is possible).

After fellowship recipients are selected and have commenced their programs at Colorado, it is intended that a second selection process will take place in order to determine which of these students are best qualified and most highly motivated to pursue the doctorate. Additionally, about halfway through the first semester of the master's program, the students' records, including background information, will be made available to the consortium units, so that these agencies might examine them in light of the organization's needs and programs. It is hoped that a match might be effected whereby the intern is matched suitably with the receiving internship site. It is possible that at any given point in time, because of the needs of the receiving

organization or because of the characteristics of the fellows, certain consortium units might have two fellows serving as interns, and other consortium units, none. However, in the long run, it is expected that each consortium agency will have had essentially equal numbers of interns.

Recruitment and Selection of Participants for Intensive Institutes

The primary criterion for determining eligibility for the short term institutes will be the extent to which the person's present or future job description requires of him evaluation and development skills that are being nurtured in the institute. A second eligibility factor will be the apparent suitability of the context in which the person is working in terms of the likelihood that the trainee would subsequently have an opportunity to exhibit skills developed during the course of the institute. The organizational atmosphere facilitative of such expression will undoubtedly vary substantially from one sending organization to the next (see Assumption 6 in Section II above). The very nature of these two criteria should suggest to the reader that the process of selecting institute participants will be relatively subjective.

To a very large extent, the intent is not to exclude from attending those persons who want to partake of institute training. However, in terms of the quality of instruction, it is apparent that some enrollment limits must be established. When a very large group of potential institute participants is identified, it is suggested that scheduling additional institutes in that general domain is an alternative preferable to increasing the number of participants to a level that is not feasible.

Of special note is the implicit assumption being made that justifies the placing of extensive resources in the institute program (see Assumption 4 in Section II). Simply stated, it is assumed that the institute as a mode of training will effectively add personnel to the ranks of the trained manpower pool in evaluation and development, at least in specialized subject areas. We feel that there presently is not an elaborate network of evaluators and developers in education. Yet we feel that the manpower pool needs to be filled and that the institute program offers at least a partial solution. Institute training should help personnel move into positions that can and will be defined differently because of the new skills of the role incumbent.

Because of the flexibility of the content of the institutes, made possible by the training package mods, it is anticipated that certain specific groups will be sought out to determine their interest in having an intensive institute tailored to their needs. Thus, for example, the regional labs might be contacted to determine their needs in development. A second institute might be planned around the needs of state department of education employees, or a third built around the demands of Title I, ESEA, evaluators in the schools. A fourth might be planned around the apparent needs of the current Ford Foundation leadership development fellows. A fifth might be constructed for the early childhood fellows taking the program at the John F. Kennedy Child Development Center, and so forth. Flexibility is exciting because it enhances the probability of establishing a close match between the instructional program and the needs of the trainees. It is further

assumed that for certain groups it will be more feasible to transport the instructional staff to a centrally-located site rather than to require all the institute participants to come great distances to Colorado.

VII. Proposed Materials Development

The development of materials is perceived in two main areas: the construction of module packages for intensive institutes; and the development of several substantial simulation activities. These two efforts are explicated in the sections that follow.

Training Materials Modules

Training material modules have already been alluded to substantially in Section IV above on the intensive training institutes. The effort to develop several training package modules is seen as one that will involve engaging on a consultant basis anywhere from fifteen to twenty professionals in the fields of education and development. A list of over twenty possible topics for these training mods was presented above in Section IV.B, and will not be repeated here. In the implementation of the program, one of the first steps will be to identify competent persons to develop these packages after first ascertaining their willingness to also serve as an instructor in an institute setting. Several persons have already been identified as excellent resource persons for this task once the program is funded.

Developers of materials in this area will be allowed considerable flexibility in terms of the format and structure that they follow. However, in the agreement with them, it will be stipulated that one of their prime responsibilities will be to develop a training package mod that consists of explicated procedures and materials to be used in training. To this extent, it is hoped that much of what is developed will have exportable characteristics, and can be used in other training

settings by other trainers. At the same time, it is felt that the same person who develops the materials should initially field test them in an actual training institute situation to increase their validity and viability.

The second major area of materials development involves simulation materials.

Simulation Materials

A considerable portion of the training of evaluators will be accomplished through the use of simulation materials. These materials represent real or hypothetical data, data collection techniques, and data analyses with which the evaluation of an educational program can be recreated. The materials presently under development are of three types: Simulation I -- the evaluation of a federal program at the national level; Simulation II -- evaluation of instructional outcomes at the state level; and Simulation III -- evaluation of a school program at the district level. Their appropriateness for graduate level training and for certain institute participants is obvious. Detailed descriptions of each of these three simulations follow.

1) Simulation I: Evaluation of a federal program at the national level.

From October, 1969, to September, 1970, the Laboratory of Educational Research, University of Colorado, was prime contractor for data analysis and reporting of the 1969 survey on compensatory education of the U. S. Office of Education. This survey was administered to a representative sample of districts and elementary schools operating Title I programs under the Elementary and Secondary Education Act (ESEA) of 1965.

The questionnaires, data analyses, data tapes, and attendant materials involved in this evaluation have been retained by the Laboratory of Educational Research. These materials will serve as the basis for a simulation of the evaluation of a federal program at the national level.

In April, 1969, a sample of the 9,236 operating programs for the disadvantaged under Title I, ESEA, was administered survey questionnaires in the national evaluation of Title I programs. Questionnaires were administered to superintendents, principals of elementary schools, teachers in grades 2, 4, and 6, and those teachers' pupils. The resulting data apply to a population of nearly 33,000 elementary schools employing over 200,000 teachers in grades 2, 4, and 6, in which approximately six million pupils were enrolled. From the questionnaire data were extracted over 400 variables descriptive of districts, schools, teachers, pupils, pupils' families, communities, the operations of compensatory education programs, and their outcomes. Data analyses resulted in over 3,000 statistical tables reporting the results of the survey.

These materials will serve as the basis of a simulation training program. Trainees will be given journalistic accounts, reports of a few local compensatory education programs, copies of questionnaires, and copies of ESEA, 1965, and its 1967 Amendment. These materials must be studied to derive general evaluative questions to be addressed in the simulation. Trainees will be encouraged to systematize the process of asking evaluative questions by devising a general framework for doing so. After evaluative questions have been stated and reviewed by the

trainer, the trainee will move to the second stage of the simulation. At this stage he will attempt to identify data sources among the more than 400 available variables which bear on answers to his evaluative questions. After so doing, the congruence between questions and data sources will be checked by the trainer to determine whether the trainee's selection of data was relevant and exhaustive. At the third stage of the simulation, the trainee will reference the data bank of over 3,000 contingency tables and extract those analyses relevant to his evaluative questions. At the fourth and final stage, the analyses will be interpreted by the trainee, and an evaluative report constructed around the trainee's evaluative questions will be written. This report will be reviewed and criticized by the trainer.

It is entirely possible, even likely, that no two trainees among dozens of trainees will produce exactly the same evaluation report. Trainees may turn up important relationships among variables in the data that have yet to be observed and recorded. Under some circumstances new analyses which appear particularly promising can be recreated from the data tapes. The simulation appears to be uniquely realistic.

2) Simulation II: Evaluation of school performance at the state level.

Since the Spring of 1967, the Laboratory of Educational Research has collaborated with the Colorado Department of Education in the development of a statewide program for the assessment of school performance. This program has been modeled on the National Assessment of Educational Progress, which is directed by the Education Commission of the States, Denver, Colorado. The program involves the assessment of pupils'

knowledges, skills, and attitudes in ten areas of the curriculum at four grade levels. To date, the program is operational in six curricular areas at four grade levels. Curriculum areas are mathematics, science, health, physical education, music, and language arts. These subject matters are assessed at grade levels 3, 6, 9, and 12.

The first stage in the development of the Colorado assessment program was the statement of behavioral objectives in ten curricular areas for grades 3, 6, 9, and 12. The objectives stated by curriculum specialists were revised and edited by measurement specialists and then administered to a sample of 1,000 teachers in a survey. The purpose of the survey was to determine which objectives would be endorsed by more than ninety percent of the teachers as those objectives they attempt to achieve in their classes. Only those objectives which received high endorsement were retained as a basis for the assessment program. Working from the agreed-upon objectives, objectives-referenced test items were devised by measurement specialists in collaboration with curriculum experts. A universe of over 500 objectives and test items resulted. In May, 1970, these items were administered to over 12,000 pupils at the four grade levels in the six curriculum areas. A matrix sampling design was employed so that the testing time for any one pupil was held below thirty minutes. The resultant data were tabulated in the form of numbers and percent correct on each item. The numbers and percent correct by item were further cross-tabulated by such biographical data on members as ethnic group membership, sex,

place of residence, home language other than English, and socio-economic status. The objectives, the test items, and the data tabluations constitute the materials for the second simulation.

In the simulation, trainees will select a subject matter and a grade level of particular interest to them at which they will work. At the first stage of the simulation, trainees will study the behaviorally-stated objectives derived through the work of curriculum specialists, measurement experts, and the survey of teachers in the field. The trainee's task at this stage will be to critique and revise these objectives. At the second stage, trainees will write objectives-referenced test items. In collaboration with the trainers, the test items and restated objectives will be evaluated. At the third stage, the trainee's test items will be compared with those test items already produced by the assessment program staff. An attempt will be made to draw a correspondence between the two sets of items. The trainee will then accept the extant items as surrogates for those he has produced. At the fourth stage of the simulation, the trainee will reference the data bank with the test items corresponding to those he produced. The trainee will interpret data analyses from the data bank and produce an evaluation report for the subject matter and grade level that he has selected. The purpose of the report will be to communicate the results to teachers, administrators, and lay persons, so that curricular changes can be made.

3) Simulation III: Evaluation of a school program at the local level.

The Worthen-Hock simulation materials were designed to introduce

the person having little or no previous evaluation experience to the general stages through which program evaluations pass, and to the things one must do to design evaluation activities. The simulations are intended to give exposure and experience with how evaluation techniques are applied in a realistic local situation, rather than to allow the simulation participant to develop marked skill in these techniques.

The materials comprise two separate simulation problems. Both are based in part on actual events. One is a local school evaluation problem concerning the mathematics program in an elementary school. The other is a district-wide evaluation problem (related to Title I and Title III projects). Although liberties have been taken with reality, an attempt has been made to adhere to reality by basing the simulations on real programs and real evaluation problems which arise in medium-size school districts.

The structure of both simulations is the same. The "in-basket technique" of presenting information is employed, as well as the use of slide-tape presentations to allow the participant to actually see and hear the action. The simulations are organized into two major sections: background, in which the participant is introduced to the situation and the character which he portrays; and tasks, five or six specific evaluation problems which the participant must attempt to solve.

How he solves the tasks is up to the participant. He is presented general information which may help him in the course of the simulation. In addition, there is other information which he may request from his "secretary" (usually the instructor). Each simulated evaluation task

is built upon the preceding task, and one task must be completed before the next is begun.

The simulation materials have been developed, tried out, and revised repeatedly over a three-year period. They are now viewed by the authors as being ready to be used in training, although of course they are still open to further revision should subsequent use show it to be necessary. Simulation III seems particularly appropriate for certain institutes that will be conducted.

It is felt that Simulations I, II, and III represent an appropriately balanced approach to evaluation via this instructional mode, and that overall they will add a reality dimension to the training. It is proposed that Simulation III, already well along in development, be completed through the use of non-federal funding. Further, it is felt that Simulations I and II can be developed to a rather polished level within the budgetary limits presented in Technical Paper Number 4.

VIII. Evaluation of the Training Programs

Evaluation of training programs at the doctoral and masters level, and training institutes, will be undertaken for two primary purposes. First, to provide feedback on operations to the project personnel; and, second, to collect data on training program effects relevant to an overall evaluation of the worth of the training procedure.

Although not elaborated upon in this section, there are two other evaluative elements that are discussed elsewhere. The Project Formative Evaluator will provide continuous feedback data to the key staff in the training programs (see Section V,). He will receive and synthesize input from trainees, instructors, internship agencies, material developers, etc. Data and judgments of a more summative nature will be provided by the independent team that will evaluate the program at least once annually (see Section IX).

Evaluation of the Doctoral Training Program for Evaluators and Developers

Professional educators, social scientists, and government planners are not of one mind as to what constitutes a worthwhile training program. It is doubtful that unanimous agreement among them can be attained on any single major issue relevant to whom the next generation of educational evaluators and developers should be, or what they should be like. Fortunately however, the problem of training inquiry related personnel in education is not totally without a few standards of excellence to which the majority of responsible observers would subscribe. A concise and thoughtful statement of one such set of standards was the product of a study group on the training of educational researchers which was established jointly by the American Educational Research Association

and Phi Delta Kappa. The study group published a set of recommendations for training educational inquiry personnel. These recommendations will serve as standards against which the accomplishments of the doctoral training programs can be measured. When observations of trainee capabilities, job responsibilities, and professional performance are held up beside the following standards, judgments of the values of the training program can result.

Recommendation 1: Parochialism in whatever guise must be abandoned. The disciplinary emphasis must be extended beyond psychology. New R&D models must be embraced within the research community. All university concern for research and education must be fostered.

The interdisciplinary emphasis of a training program can be evaluated through the observation of the rate of which formal course work is taken in disciplines outside education. The record of exposure of trainees to scholars in the social sciences, the observation of opportunities provided trainees to become familiar with published writing in other fields, etc. The effectiveness of exposure to a range of disciplines can be evaluated partially through the use of paper and pencil techniques, measuring knowledge of leading figures, research trends, and major findings from such disciplines as sociology, economics, political science, and psychology.

Recommendation 2: New and vigorous programs of student recruitment must be undertaken to change the character of the student body interested in research in education. Possibilities for careers in educational inquiry must be identified for bright undergraduates in many fields and these students must be encouraged and supported in their pursuit of these career lines.

Such data as the following are relevant to the evaluation of doctoral training programs in relation to this standard.

a. Test score data on trainee aptitude.

b. Biographical data descriptive of trainee educational history.

c. Data on geographic distribution of trainees.

d. Information on recruitment techniques and program publicity.

Recommendation 3: Extensive experimentation should be undertaken in institutional arrangements and program content for the training of researchers in education.

Performance in relation to this standard can perhaps best be measured through the examination of a project history in which attempts at innovation in training programs are recorded.

Recommendation 4: The program content should emphasize opportunities for early production on the part of the student, continuous involvement with productive and successful researchers, and field experience in the conduct of research, evaluation and development.

Data relevant to the performance of training programs in these respects include records of trainees' writings, both published and unpublished, time and effort reports indicating settings in which training activities are conducted, and persons under whose direction training is carried out.

Recommendation 5: The uniqueness of scholarship in education should be recognized and accounted for in planning programs for research in education. This will mean setting the stage at the graduate level so that subsequent collegial relationships can develop which will allow the researcher in education to bridge his involvement in a social process field and his commitment to a discipline of study.

Such data as bear on the above standard are difficult and costly to collect, but nonetheless, are critical to the determination of the content of trainee programs. Perhaps retrospective histories of trainees' professional growth written some years after participation in training programs are the most pertinent data that address the above concerns.

The Laboratory of Educational Research has had extensive experience in the evaluation of doctoral training program in research evaluation and development. Since 1966, the Laboratory has conducted a graduate level training program supported under Title IV of ESEA, 1965. The program training operation and outcomes have been evaluated annually. Extensive instrumentation, interview techniques, and report outlines have been developed over the years in connection with this annual evaluation. A copy of a recent evaluation report of the Title IV doctoral training program conducted by the Laboratory of Educational Research is appended to this proposal as Appendix B.

Evaluation of the Masters Training Program for Evaluators and Developers

On the following pages an outline is given of the salient questions that will be answered in order to evaluate the masters degree programs.

Outline of Evaluation Plan for

Master's Program in Evaluation and Development

Questions About Instrumentalities	Questions About Outcomes
1. Publicity a. Who hears about the program and who doesn't? How do they hear about it? b. How much "lead-time" must be given to participants? c. What information should one communicate in publicity? 2. Participant Selection a. What <u>logical</u> criteria (characteristics of participants) are related to the objectives of the program? b. What empirical criteria (age, undergraduate G.P.A., etc.) relate to "success" in the program? c. What is an optimal number of participants for such a program (10, 20, 30, or more)? 3. Pre-arrival Preparation of Participants a. How does previous academic course work relate to performance in the program? b. Can performance in the program be facilitated by requiring participants to study books on a reading list before arriving? (Will participants respond to request for directed self-study before arriving)? c. Do participants arrive with realistic expectations?	1. Attainment of Subject Matter Objectives a. How is course performance on a mastery test related to "entry behavior" (pre-arrival preparation)? b. Is there retention of subject-matter mastery one and two years after the program? (Is performance superior to a "control group" on mastery tests after one and two years)? 2. Attainment of Affective Objectives a. Are attitudes toward evaluation and development changed? 3. Long-range Outcomes (One or Two Years Later) of Participation a. Cognitive: Do participants <u>achieve</u> at a higher level than a nonparticipating "control-group" one and two years after the program? b. Affective: Are changes in attitudes and values maintained? c. Professional: 1. Do participants assume roles for which they were trained? 2. Do they seek academic work beyond the M.S. degree at a rate higher than that for a "control group"?

Questions About Instrumentalities	
4. Conduct and Content of Course Work a. Is the content of the course work relevant to the objectives of the program in the opinion of evaluation and development experts, educationists, the participants and staff? b. In what ways do the courses suffer from poor organization and implementation in the opinion of the participants, staff, and outside observers? c. What factors facilitate or impede training in the simulation exercises? 5. Other Instructional Activities a. What types of lecture and small group activity facilitate attaining the objectives of the program? b. How often and when should auxiliary instructional activities be scheduled? c. Should "tutoring" by doctoral students be provided to the participants? 6. Support and Housing of Participants a. Is financial support adequate? b. Should an effort be made to house participants close together?	

By no means do the questions recorded above exhaust all possibilities. Many important questions cannot now be anticipated; thus the necessity of a flexible evaluation plan responsive to the evolving needs of the program is apparent.

Evaluation data which bear on the decisions identified above fall into three classes: antecedents (conditions existing at the outset of the program), transactions (the instructional and other activities which comprise the program), and outcomes (the results of the program).

A. Antecedents.

1. Publicity.

a. Scope.

b. Effectiveness of communication.

c. Timing.

2. Participants.

a. Biographical and professional data.

1. Age.

2. Sex.

3. Undergraduate training.

4. Nature of employment.

b. Preparation for this program (entry behavior).

3. Objectives.

a. Formal courses.

b. Teaching practicum.

c. Informal activities.

B. Transactions.

1. Course Work.

a. Subject matter.

b. Organization.

c. Evaluation (testing and grading): complaints and satisfactions.

2. Other Instructional Activities.

a. Seminars.

b. Small group discussions.

c. Tutorial sessions.

d. Simulation exercises.

C. Outcomes.

1. Participants.

a. Subject-matter mastery.

b. Affective outcomes.

c. Satisfactions and complaints.

d. Financial costs (direct and indirect).

e. Professional (job mobility, promotion, further academic training, etc.).

2. Staff.

a. Satisfactions and complaints.

b. Costs (financial and time).

The following table, Table 5, shows when and how the evaluation data identified above will be gathered. (In the following table, data are identified by the outline designations in the preceding section. Evaluation activities are identified by letters which are elaborated after the table.)

Table 5 Data collection calendar

Data	June '71	July '71	August '71	September '71	October '71	November '71	December '71	January '72	February '72	March '72	April '72	May '72	June '72	July '72	August '72	. . .	September '73	September '74
A. Antecedents 1. a-c	A																	
2. a-b				B,C														
3. a-c			D															
B. Transactions 1. a-c						E_1		E_2			E_3		E_4		E_5			
1 & 2								F_1					F_2		F_3			
1 & 2				G_1	G_2	...								...	G_{12}			
C. Outcomes 1. a								H_1					H_2		H_3		H_4	H_5
1. b				I_1											I_2		I_3	I_4
1. c															J_1		J_2	
1. d															K			
1. e																	L_1	L_2
2. a															M			
2. b															N			

A. The scope, timing, and effectiveness of publicity will be assessed by means of questionnaires or telephone surveys of a random sample of agencies which are potential suppliers of trainees. In addition, participants will also be questioned about where they heard of the program, how, when, whether publicity was too late or appropriately timed, etc.

B. Biographical and professional data on the participants will be gathered by means of a questionnaire. These data will serve both as a description of the group of participants and as baseline data against which to measure change over time (see below).

C. Mastery tests over course content will be administered at the outset to be used as feedback to instructional staff on the entering level of the participants.

D. Staff will be asked to record--in behavioral terms--the objectives of their instruction, both formal and informal.

$E_1 - E_5$. Course work will be evaluated by the participants by means of course evaluation inventory (see Attachment 1 in Appendix C) to which several "ad hoc" questions will be added as the need arises.

$F_1 - F_3$. A questionnaire will be filled out by the staff of the program at the end of the first semester (January 1972), the second semester (June 1972), and the end of the summer session (August 1972).

$G_1 - G_{12}$. In September of 1971, one or more of the participants will be designated a "participant observer" whose job it will be to serve as a liaison between the participants and staff, to provide direction to the staff based on his perception of participants' reaction, etc. (i.e., a junior ombudsman).

$H_1 - H_5$. Subject matter mastery will be tested routinely following each course. In addition, a follow-up test of retention of subject-matter will be administered in September 1973 and September 1974.

$I_1 - I_4$. Attitudes toward educational evaluation will be assessed before and after the program and again in September 1973 and 1974. Examples of attitude inventories appear as Attachment II in Appendix C.

$J_1 - J_2$. Participants will fill out a critique sheet on the objectives, operations, and outcomes of the program. See Attachment III in Appendix C; note that it would also be appropriate for the short-term institutes if minor modifications were made.

K. Participants will be asked to indicate (anonymously) the costs (both direct and indirect) which were incurred in the course of the year.

$L_1 - L_2$. In September of 1973 and 1974, a questionnaire on professional activity will be administered.

M and N. Questionnaires assessing satisfactions and complaints and costs will be administered to the staff of the program.

To the extent desired and feasible, measurements A through N can be repeated with each subsequent group of entering trainees.

Evaluation of Short-Term Training Programs for Evaluators and Developers

Carefully planned and executed evaluation of training institutes is perhaps even more crucial than evaluation of extended masters and doctoral programs. Because of their short duration, feedback to project personnel on operations must be rapid and quickly acted upon. An outline of an evaluation plan for short-term training programs is not herein presented, but would be greatly similar to the plan presented in the previous section for the evaluation of the masters' program.

Moreover, because of the possibility that training will not bring about permanent nor massive changes in skills and behaviors, the evaluations of outcomes of the intensive training institutes is of utmost importance. Follow-up testing and interviewing can be easily accomplished to measure extended program effects. It can be noted that short duration institutes have had some measurable long-range effects (see Appendix D on the follow-up of the 1967 AERA presession on experimental design). Similar evaluative efforts will be conducted for these training institutes.

The reader should recall that an experiment is planned to determine the relative effectiveness of different follow-up models. Short-term institute trainees will be randomly assigned to one of four treatment groups:

1) Traditional, i.e., no follow-up.
2) Phone and letter follow-up.
3) Treatment 2 plus periodic follow-up sessions on campus.
4) Treatments 2 and 3 plus supervision and assistance on-the-job at the trainee's home base (primarily by the Associate Director for the Intensive Institutes assisted by selected graduate students).

Measurements will be made of the trainees attitudes, skill levels, and extent of using each skill.

Evaluation of Materials Developed

For the most part, the simulations and the training material mods developed will be evaluated on the basis of user satisfaction as well as changes in user skill level. As modifications are made in the materials after formative evaluation data is incorporated, user satisfaction and skill level will again be ascertained to insure that the modifications result in changes in a positive direction.

IX. Administering, Monitoring, and Determining the Cost Effectiveness of The Training Program

Presented below is information bearing on administering and monitoring the training programs, as well as on determining their cost effectiveness. Also included is a brief consideration of the possibilities for other means of supporting such a program once federal funding is reduced or withdrawn.

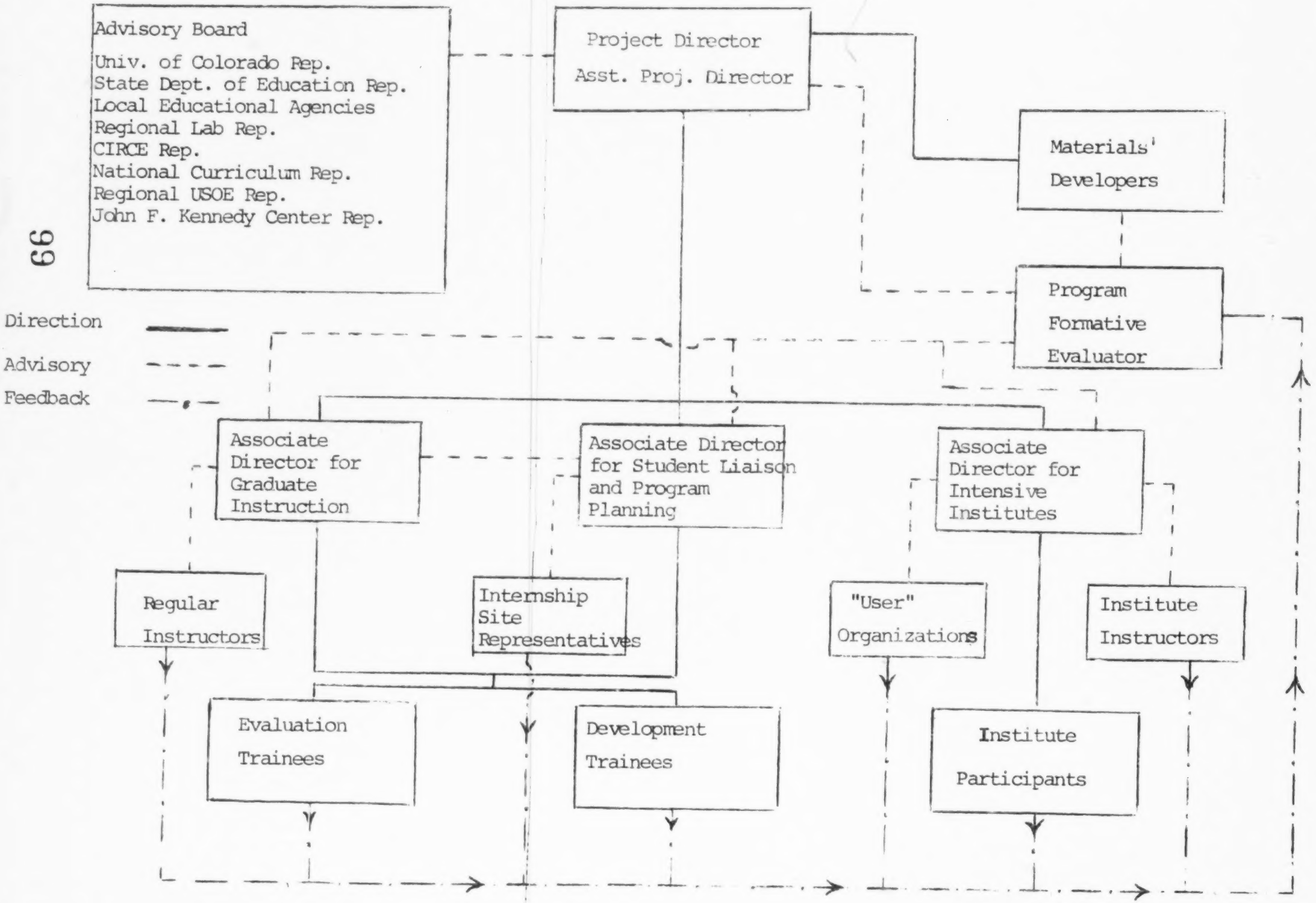
Administration of the Training Program

The University of Colorado will serve as the prime contactor for the training consortium. As such, fiscal and accounting responsibility will rest with one agency. Suitable accounting procedures have long been practiced by the Office of Research Services for the University. Elaborate precautions will be taken to insure that the administrative paper work is kept at a minimum. Project administration would be mainly the responsibility of the Project Director. Assisting him in this function would be the Assistant Director for the training program.

Within the program itself, the operations of an administrative nature are in part presented graphically on the next page. Depicted in the diagram are patterns of relationships that are likely to develop within the training program.

As can be seen in the chart, the Project Director receives advice and counsel from the Advisory Board, which is made up of representatives from the consortium units. He also receives recommendations from the Program Formative Evaluator. The Project Director deals directly

PATTERNS OF RELATIONSHIPS IN THE TRAINING CONSORTIUM.



with the materials' developers, and also with the three Associate Directors. The Associate Directors receive advice from the various instructors and/or organizations with which they have the most immediate relationships, as well as from the Program Formative Evaluator. Further, it might be noted that each of the instructors and each of the trainees, as well as the internship sites and user organizations have a direct line, so to speak, to the Program Formative Evaluator, so that they might provide him with feedback relative to their perceptions of the multiple training programs.

Monitoring the Training Program

Monitoring of the program will take place at four levels. At one level, the Director and the Associate Directors will be responsible for ascertaining the effectiveness of day-to-day operations and for determining the congruence between intended transactions and those which actually take place. At a second level, the Program Formative Evaluator will serve an important monitoring function. As each of the trainees and each of the consortium organizations has immediate access to him, he is in an excellent position to perform a valuable monitoring function. A third monitoring device is built into the evaluation plan mentioned earlier. One or more of the participants will have been designated as a participant observer. His job will be to serve as a liaison between the participants and the staff, and to record, as accurately as possible, his perceptions of the participants' reactions to the training, and other relevant data.

The fourth level is somewhat more intense, impartial, and objective. In Technical Paper Number 4, it can be observed that an outside evaluation team is provided for in the budget (also see Section VIII).

with the materials' developers, and also with the three Associate Directors. The Associate Directors receive advice from the various instructors and/or organizations with which they have the most immediate relationships, as well as from the Program Formative Evaluator. Further, it might be noted that each of the instructors and each of the trainees, as well as the internship sites and user organizations have a direct line, so to speak, to the Program Formative Evaluator, so that they might provide him with feedback relative to their perceptions of the multiple training programs.

Monitoring the Training Program

Monitoring of the program will take place at four levels. At one level, the Director and the Associate Directors will be responsible for ascertaining the effectiveness of day-to-day operations and for determining the congruence between intended transactions and those which actually take place. At a second level, the Program Formative Evaluator will serve an important monitoring function. As each of the trainees and each of the consortium organizations has immediate access to him, he is in an excellent position to perform a valuable monitoring function. A third monitoring device is built into the evaluation plan mentioned earlier. One or more of the participants will have been designated as a participant observer. His job will be to serve as a liaison between the participants and the staff, and to record, as accurately as possible, his perceptions of the participants' reactions to the training, and other relevant data.

The fourth level is somewhat more intense, impartial, and objective. In Technical Paper Number 4, it can be observed that an outside evaluation team is provided for in the budget (also see Section VIII).

This team is to be composed of several persons representing different orientations, and also representatives from outside of education. In at least one annual visit, the team will be charged with critically examining and evaluating the operations of the entire training program, and providing direct information in this regard to the Advisory Board, the Director, the Assistant Director, the Formative Evaluator, and the Associate Directors. The primary purpose of this evaluation will be program improvement, although reports of this team will be retained and will serve as an important outside, summative-type of evaluation that is considered valuable. By means of these four levels of monitoring, it is expected that relevant data can be fed back into the program to allow improvement.

Determining the Cost Effectiveness of the Program

The budget breakdowns in Technical Paper Number 4 give some indication of the relative resources that are being expended on each of the various elements of the training program. With relatively simple mathematics, it can be determined from those breakdowns how much it is costing to train each different type of participant; that is, a different cost figure can be arrived at for a doctoral student as compared with a master's student, as compared with an institute trainee.

More important and variable, at least in the case of the institute participants, is the relative cost of providing the various degrees of follow-up. It is expected that, by keeping careful time records during the first year of operation, relative cost data will be available on resource expenditures for each of the four possible follow-up

models. Given this cost data, and the resultant effectiveness of the various follow-up models on trainees, it should be possible to arrive at a feasibility and cost effectiveness estimate for each of the four follow-up alternatives.

Continuation of the Program

It is hoped and expected that federal funds for the support of this consortium will be available for several years. If this is the case, it will allow the consortium training program to become firmly established as a viable and productive enterprise. It can be noted in the fourth technical paper in this series that the general intent is to keep the total budget at about the same level as the second year. Changes from year to year are minor except for less being expended on materials development each year and slightly more being expended on training. Additionally, it can be noted that percentages of time for administration of the program reduce slightly after the first year, as it is felt that certain of the operations will become less demanding, and also that the administrative responsibilities with regard to the development of materials will be reduced.

It is always appropriate at the outset of a new program to look at the possibilities for its continuation under various alternative funding provisions. It is obvious that the federal government cannot continue to support operations forever, and this is as it should be. The feeling here is that there could develop institutional and consortium support for continuation of this type of training after federal funding is withdrawn. It is considered possible that with a nucleus for training thus established it will be possible to gradually reconstitute

models. Given this cost data, and the resultant effectiveness of the various follow-up models on trainees, it should be possible to arrive at a feasibility and cost effectiveness estimate for each of the four follow-up alternatives.

Continuation of the Program

It is hoped and expected that federal funds for the support of this consortium will be available for several years. If this is the case, it will allow the consortium training program to become firmly established as a viable and productive enterprise. It can be noted in the fourth technical paper in this series that the general intent is to keep the total budget at about the same level as the second year. Changes from year to year are minor except for less being expended on materials development each year and slightly more being expended on training. Additionally, it can be noted that percentages of time for administration of the program reduce slightly after the first year, as it is felt that certain of the operations will become less demanding, and also that the administrative responsibilities with regard to the development of materials will be reduced.

It is always appropriate at the outset of a new program to look at the possibilities for its continuation under various alternative funding provisions. It is obvious that the federal government cannot continue to support operations forever, and this is as it should be. The feeling here is that there could develop institutional and consortium support for continuation of this type of training after federal funding is withdrawn. It is considered possible that with a nucleus for training thus established it will be possible to gradually reconstitute

the training program as the Evaluation and Development Training Unit of the Laboratory of Educational Research. It is felt that if the training has been successful and if the persons so trained become increasingly effective in their occupational responsibilities, then support for such an operation could come largely from contracts various school districts, organizations needing evaluation design assistance, groups that desire certain types of institute training to be conducted, and the like. Although it may not be possible to continue such a center at the same level of funding, it is expected that enough funds could be secured to continue the program at a viable, productive, and meaningful level.